END
DATE
FILMED
10-79
DDC

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

CENTER FOR MARITIME STUDIES

WEBB INSTITUTE OF

NAVAL ARCHITECTURE

LEVEL

LONG-TERM DISTRIBUTIONS OF
STOCHASTIC PROCESSES
WITH
APPLICATION TO SHIP
HULL STRESS STATISTICS

GLEN COVE
NEW YORK
11542

by
Bruce H. Stephan
and
Robert B. Zubaly

1889

December 1978

CENTER FOR MARITIME STUDIES
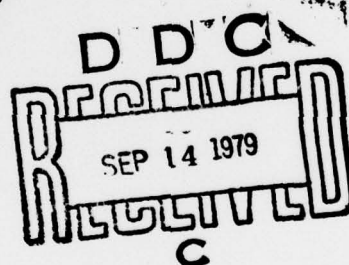
WEBB INSTITUTE OF NAVAL ARCHITECTURE

CRESCENT BEACH ROAD   GLEN COVE. NEW YORK 11542

LONG-TERM DISTRIBUTIONS OF

STOCHASTIC PROCESSES

WITH

APPLICATION TO SHIP

HULL STRESS STATISTICS,

Final Rept.,

by

Bruce H. Stephan

and

Robert B. Zubaly

H11369

December 1978

71 p.

DDC RECEIVED SEP 14 1979 C

79 09 13 017

411 369

ABSTRACT

The mathematical properties of the following probability
distributions and their moment generating functions are
derived:

Weibull distribution

Rayleigh distribution

Exponential distribution

Normal distribution

Voznesensky distribution

Generalized Rayleigh distribution

The relative merits of applying these distributions to
problems in ship responses to the sea, which is described
as a stationary stochastic process, are discussed.

In Part 2, the most promising long-term distributions
derived from the above survey are applied to ship bending
stress data from four ships. It is concluded that a
numerical solution of long-term distributions, using either
Weibull or Normal distributions of the short-term Rayleigh
parameters (classified by weather groups) is better than
any explicit function.

TABLE OF CONTENTS

Part 1

Mathematical Properties

Part 1
Mathematical Properties

1. Introduction

Researchers of the past decade have studied the waves of the ocean's surface and the accompanying wave-induced stresses in a ship's hull by treating these quantities as continuous random functions which may be represented by stationary stochastic processes. The wave height, in terms of displacement from the mean, may at any instant be viewed as the net displacement resulting from the sum of a large number of small displacements, each of which is the result of a single wave originating over a wide area and which is independent of the others. The central limit theorem of probability theory guarantees that the sum of a large number of independent random variables may be treated as if it were itself normally distributed. Thus the process governing wave height may be treated as normally distributed. Since a ship's response is linearly related to the wave height, then the response may also be viewed as the sum of a large number of independent random variables and hence also treated as if it were normal. Thus the stationary stochastic process of interest is assumed to be normal.

The ship designer is concerned with many practical questions about this normally distributed stochastic process. If the displacement from the mean is normally distributed, what is the distribution of the peak-to-trough values in many reversals? What are the chances of exceeding a particular stress level in one reversal, or in an operating period, or in a ship's lifetime? What are the chances that the maximum stress in any period does not exceed a certain level? If the questions depend on weather or sea state conditions, what are the answers to the above questions with and without regard to weather and sea state? What are the chances of encountering different weather and sea conditions? And, in short, to what value should one design his structure with a pre-determined risk factor?

These questions involve the distributions of many random variables which are related to the normally distributed stochastic process. Numerous data in the form of twenty-minute wave or stress records have been laboriously collected in an attempt to recognize which distributions are at work, and what are the values of any parameters present. It is the purpose of this part of the report to catalog the common and not-so-common probability distributions which researchers have used and to derive their significant properties. Hopefully, this will contribute to the more intelligent use of recorded data.

2. Basic Probability Concepts[*]

Before we embark on a discussion of the different distributions, we shall briefly discuss some of the elementary probability ideas which are used in the comparative discussion which follows.

A continuous random variable X is described by two functions f(x) and F(x) which are respectively called the probability density function (pdf), and the cumulative distribution function (cdf) for X. F(x) is the probability that the random variable X takes on a value less than or equal to x, while $f(x) = \frac{dF}{dx}$. Both functions are loosely referred to as "the distribution of X" and f(x) may be loosely interpreted as the probability that X takes on a value in some "small" interval about x. The probability that X falls in some interval [a,b] is

$$P(a \leq X \leq b) = \int_a^b f(t)\, dt$$

Thus, F(x) is the area under the density curve up to value x.
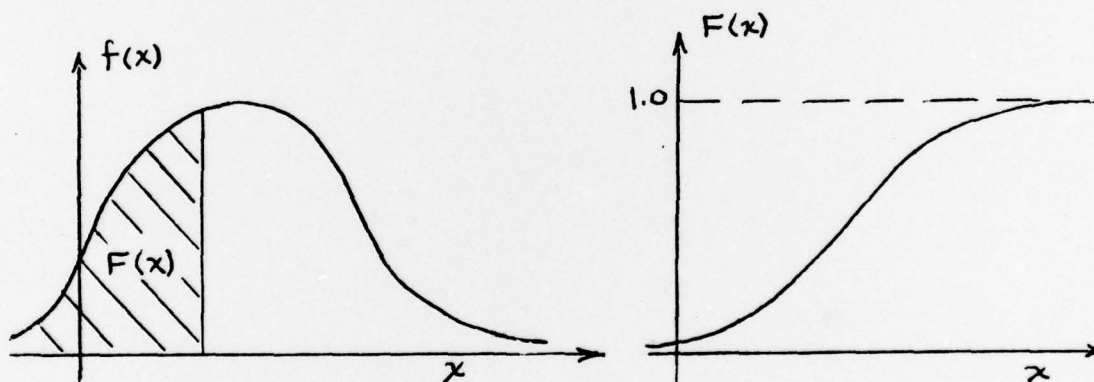
The distributions of random variables can be characterized by their moments. The nth moment of a random variable is defined as

$$E(X^n) = \int_{-\infty}^{\infty} x^n f(x)\, dx,$$

and the first two moments are most frequently used in the form of the mean $\mu = E(X)$ and the variance $\sigma^2 = E(X^2) - E(X)^2$. The more moments, starting with the mean, which two different distributions have in common, the more similarly the two random variables behave.

The moments of a random variable are exceedingly useful in estimating any unknown parameters from a statistical sample. Suppose one has a sample $x_1, x_2 \ldots$

---

[*] A more detailed discussion of the basic concepts may be found in Freund (1), Feller (2), Meyer (3). See references listed on pp. 40-41.

x_k from a population where the unknown distribution is thought to have a certain form but with unknown parameters. Then

$$\frac{x_1^n + x_2^n + \ldots + x_k^n}{k}$$

may be taken as an estimate of $E(x^n)$. If theoretical expressions for $E(x^n)$ have been developed in terms of the unknown parameters, then these estimates of $E(x^n)$ give equations which may be solved for the parameters.

The moment generating function (MGF) of a random variable is a frequently used device to develop the moments of a random variable. If X is a random variable with pdf $f(x)$ then the MGF, $M_x(\theta)$ is defined as

$$M_x(\theta) = E(e^{x\theta}) = \int_{-\infty}^{\infty} e^{x\theta} f(x)\, dx$$

Since
$$e^{x\theta} = \sum_{n=0}^{\infty} \frac{x^n}{n!} \theta^n \ , \quad M_x(\theta) = \sum_{n=1}^{\infty} \frac{\theta^n}{n!} E(x^n) \quad \text{and hence}$$

$$\left. \frac{d^n M_x(\theta)}{d\theta^n} \right|_{\theta=0} = E(x^n) \ .$$

Several other useful properties of moment generating functions which greatly simplify the comparison of random variables are listed below.

1) There is a one-to-one relation between moment generating and distributions of random variables. That is if X and Y are random variables whose MGF's $M_X(\theta)$ and $M_Y(\theta)$ both exist and are equal, then X and Y have the same distribution.

2) If the density of random variable X has a parameter λ and as $\lambda \to \lambda_0$, $M_X(\theta) \to M_Y(\theta)$, then the distribution of X may be treated as the distribution of Y for λ close to λ_0.

3) If X has MGF $M_X(\theta)$ and $Y = \frac{X + a}{b}$, then Y has MGF $M_Y(\theta) = e^{a/b\, \theta} M_X(\frac{\theta}{b})$.

4) If $X_1, X_2, \ldots, X_n$ are independent random variables with respective MGF's $M_{X_1}(\theta), \ldots, M_{X_n}(\theta)$ and $Y = X_1 + X_2 + \ldots + X_n$, the MGF $M_Y(\theta) = M_{X_1}(\theta) M_{X_2}(\theta) \ldots\ldots M_{X_n}(\theta)$.

We require an additional property of probability density functions. If X and Y are two random variables with respective pdf's $f(x)$ and $g(y)$ which are related in a one-to-one fashion through a steadily increasing or decreasing function $x = h(y)$ then
$$g(y) = f(h(y)) \left| \frac{dx}{dy} \right|$$

-3-

3. The Weibull Distribution

The Weibull distribution has been used of late by a large number of researchers (Nordenstrom (4), Mansour (5), Hoffman and Karst (6) Voznessensky (7)), in dealing with engineering problems related to the ocean's surface and may be called "currently in fashion". The high flexibility of its two or three parameters allow for excellent curve fitting to a wide variety of data, and many earlier used distributions may be made special cases of the Weibull distribution.

A random variable is said to have a Weibull distribution if its pdf has the form

$$f(x) = \begin{cases} \alpha \beta x^{\beta-1} e^{-\alpha x^\beta} & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad \text{where } \alpha > 0, \ \beta > 0$$

The cumulative distribution function is

$$F(x) = \begin{cases} 1 - e^{-\alpha x^\beta} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

This is the two-parameter form which appears in most references. There is a three-parameter form of this distribution which shifts the left end of the density curve further leftward and allows for negative values of the random variable X. This alternate form of the Weibull distribution may be expressed as density

$$f(x) = \begin{cases} \gamma \left(\frac{x-a}{b}\right)^{\gamma-1} e^{-\left(\frac{x-a}{b}\right)^\gamma} & x \geq a \\ 0 & x < a \end{cases}$$

and cumulative distribution

$$F(x) = \begin{cases} 1 - e^{-\left(\frac{x-a}{b}\right)^\gamma} & x \geq a \\ 0 & x < a \end{cases}$$

(Note: Some references prefer to write b as V-a)

We shall calculate the mean and variance of this random variable from its moment generating function which will be calculated first for the two-parameter form and then converted to the three-parameter form by property 3 of moment generating functions.

Now

$$M_X(\theta) = E(e^{x\theta}) = \int_{-\infty}^{\infty} e^{x\theta} f(x) \, dx = \int_{-\infty}^{\infty} \alpha \beta x^{\beta-1} e^{-\alpha x^\beta} e^{x\theta} \, dx \ ,$$

and letting $y = \alpha x^\beta$ we have

$$M_X(\theta) = \int_0^\infty e^{\theta \left(\frac{y}{\alpha}\right)^{1/\beta}} e^{-y} \, dy \ .$$

If $e^{(\frac{y}{\alpha})^{1/\beta} \theta}$ is expanded as $\sum\limits_{n=0}^{\infty} \frac{\theta^n}{n!} (\frac{y}{\alpha})^{n/\beta}$, we have

$$M_X(\theta) = \sum_{n=0}^{\infty} \alpha^{-n/\beta} \frac{\theta^n}{n!} \int_0^{\infty} e^{-y} y^{n/\beta} dy .$$

The integral $\int_0^{\infty} y^{s-1} e^{-y} dy$ is denoted by $\Gamma(s)$ and is the well known gamma function. It is a generalization of $s!$ ($\Gamma(s+1) = s!$ for s an interger) and its properties are outlined in the appendix.

Thus the MGF for the two-parameter form of the Weibull distribution is

$$M_X(\theta) = \sum_{n=0}^{\infty} \frac{\theta^n}{n!} \alpha^{-n/\beta} \Gamma(1 + \frac{n}{\beta})$$

and the nth moments are

$$E(X^n) = \alpha^{-n/\beta} \Gamma(1 + \frac{n}{\beta}) .$$

If X has the above 3-parameter form, then $z = x-a$ is a random variable fitting the two-parameter Weibull density with $\beta = \gamma$ and $\alpha = \frac{(1)^{\beta}}{b^{-\beta}}$. Thus

$$M_z(\theta) = \sum_{n=0}^{\infty} b^n \Gamma(1 + \frac{n}{\gamma}) \frac{\theta^n}{n!} .$$

Since $x = z+a$, property 3 of moment generating functions gives

$$M_X(\theta) = e^{a\theta} \sum_{n=0}^{\infty} b^n \Gamma(1 + \frac{n}{\gamma}) \frac{\theta^n}{n!}$$

as the MGF of the three-parameter Weibull distribution.

Since the a in the three-parameter form is merely a translation, we shall restrict our investigation of properties to the two-parameter form unless forced to use the three-parameter case. The mean of the three-parameter form is merely the mean of the two-parameter form translated by a and the variances are the same.

First

$$\mu = \alpha^{-\frac{1}{\beta}} \; \Gamma \left(1 + \frac{1}{\beta}\right)$$

and

$$\sigma^2 = E(X^2) - E(X)^2$$

$$= \alpha^{-\frac{2}{\beta}} \; \left[\; \Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right) \;\right]$$

Observe that the parameter α may be eliminated, and the simple relation

$$\frac{\sigma^2 + \mu^2}{\mu^2} = \frac{\Gamma\left(1 + \frac{2}{\beta}\right)}{\Gamma^2\left(1 + \frac{1}{\beta}\right)} \equiv R(\beta)$$

established between μ and σ and the β parameter. Since μ and σ may be estimated from a statistical sample, the β parameter may be easily estimated. Unfortunately, the β may not be explicitly solved for and one must resort either to a numerical solution to the equation or to examining the graph of $R(\beta)$ in Figure 1. Note also that as β increases, $R(\beta)$ asymptotically approaches 1 and hence the larger μ becomes, the larger β will be.

The MGF $M_X(\theta)$ may be written as

$$M_X(\theta) = \sum_{n=0}^{\infty} \left(\frac{\theta \alpha^{-1/\beta}}{n!}\right)^n \Gamma\left(1 + \frac{n}{\beta}\right)$$

or

$$M_X(\theta) = M_Y\left(\theta \, \alpha^{-1/\beta}\right) = M_{\alpha^{-1/\beta} Y}(\theta)$$

where Y is Weibull with $\alpha = 1$ and parameter β. Thus $X = \alpha^{-1/\beta} Y$, and the α appears to be playing the role of a change of scale in the Weibull distribution. The β parameter seems then to characterize the type of distribution which must first be determined from data. Once the β is determined then the α may be estimated either from

$$\mu = \alpha^{-1/\beta} \; \Gamma\left(1 + \frac{1}{\beta}\right)$$

or from a maximum likelihood estimate with confidence intervals as will be discussed later.

-6-

The above discussion implies that a scalar multiple of a Weibull variable is still Weibull distributed. This is so, for if $Y = \lambda X$ where X has pdf

$$f(x) = \begin{cases} \alpha \beta x^{\beta-1} e^{-\alpha x^\beta} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

then Y has pdf

$$g(y) = \begin{cases} \alpha \beta \left(\frac{y}{\lambda}\right)^{\beta-1} e^{-\alpha \left(\frac{y}{\lambda}\right)^\beta} \frac{1}{\lambda} & y \geq 0 \\ 0 & y < 0 \end{cases}$$

or

$$g(y) = \begin{cases} \bar{\alpha} \beta y^{\beta-1} e^{-\bar{\alpha} y^\beta} & y \geq 0 \\ 0 & y < 0 \end{cases}$$

where $\bar{\alpha} = \alpha/\lambda^\beta$. This means that all Weibull random variables are linearly related to a standardized Weibull variable with $\alpha = 1$. Hence tables for "standardized variables" at different β values could be prepared which would greatly simplify use of this distribution.

4. Maximum Likelihood Estimate and Confidence Intervals

If $X_1, X_2, \ldots X_n$ are n independent Weibull random variables with the same fixed β parameter and identical distributions

$$f(x) = \alpha \beta x^{\beta-1} e^{-\alpha x^{\beta}}$$

then the liklihood function $L(\alpha)$ for a set of values $x_1, x_2, \ldots x_n$ is defined as the probability of $X_1 = x_1, \ldots, X_n = x_n$. Thus

$$L(\alpha) = \alpha \beta x_1^{\beta-1} e^{-\alpha x_1^{\beta}} \ldots \alpha \beta x_n^{\beta-1} e^{-\alpha x_n^{\beta}}$$

$$= \alpha^n \beta^n (x_1 x_2 \ldots x_n)^{\beta-1} e^{-\alpha \sum_{i=1}^{n} x_i^{\beta}}$$

The value of α which maximizes $L(\alpha)$ then is taken as the "best" estimate α for this sample and is called the maximum likelihood estimate.

(Note: The maximum likelihood estimate may or may not agree with the value obtained through use of α's relation to the mean. These are two different estimates which reflect two different criteria of "best" estimate. They may frequently agree but in general need not. The advantage of the maximum likelihood estimate in this case is that a confidence interval may be developed for the true α value of the sampled population rather than a simple value with no error bounds).

$L(\alpha)$ may be maximized by setting $\frac{dL}{d\alpha} = 0$. This is awkward, however. Since

$$\frac{d \ln L(\alpha)}{d\alpha} = \frac{1}{L(\alpha)} \frac{dL}{d\alpha}$$

we shall solve

$$\frac{d \ln L(\alpha)}{d\alpha} = 0$$

instead.

$$\ln L(\alpha) = n \ln \alpha + n \ln \beta + (\beta-1) \sum_{i=1}^{n} \ln x_i - \alpha \sum_{i=1}^{n} x_i^{\beta}$$

$$\frac{d \ln L(\alpha)}{d\alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} x_i^{\beta}$$

and thus

$$\alpha^{-1} = \frac{1}{n} \sum_{i=1}^{n} x_i^{\beta}$$

is the "best" estimate of α.

A confidence interval for α, or equivalently $1/\alpha$ may be developed as follows. Consider the random variable

$$U = \frac{1}{n} \sum_{i=1}^{n} X_i^{\beta}$$

where X_i are independent identically distributed Weibull random variables. U is itself a random variable with a distribution yet to be determined.

First, if X has pdf

$$f(x) = \begin{cases} \alpha x^{\beta-1} e^{-\alpha x^{\beta}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$Y = x^{\beta}$ has pdf

$$g(y) = \begin{cases} \alpha \beta y^{\beta-1/\beta} e^{-\alpha y} \left| \frac{dx}{dy} \right| & y \geq 0 \\ 0 & y < 0 \end{cases}$$

or

$$g(y) = \begin{cases} \alpha \beta y^{\beta-1/\beta} e^{-\alpha y} \frac{1}{\beta} y^{1-\beta/\beta} & y \geq 0 \\ 0 & y < 0 \end{cases}$$

or

$$g(y) = \begin{cases} \alpha e^{-\alpha y} & y \geq 0 \\ 0 & y < 0 \end{cases}$$

Thus, X^{β} is Weibull distributed with parameter $\beta = 1$ and α as in the sampled population. Hence, Y has MGF

$$M_y(\theta) = \sum_{n=0}^{\infty} \left(\frac{\theta}{\alpha}\right)^n \frac{1}{n!} \Gamma(n+1)$$

$$= \sum_{n=0}^{\infty} \left(\frac{\theta}{\alpha}\right)^n = \left(1 - \frac{\theta}{\alpha}\right)^{-1}$$

By property 4) of moment generating functions, U then has MGF

$$M_{nU}(\theta) = \left(1 - \frac{\theta}{\alpha}\right)^{-n}$$

or

$$M_{\alpha 2nU}(\theta) = (1 - 2\theta)^{-n}$$

by property 3).

The random variable W with pdf

$$f(w) = \begin{cases} \dfrac{1}{2^{m/2}\,\Gamma(m/2)}\ w^{\frac{m-2}{2}}\ e^{-\frac{w}{2}} & w > 0 \\[2mm] 0 & w \le 0 \end{cases}$$

is called chi-squared distributed with m degrees of freedom. It's MGF is

$$M_w(\theta) = (1 - 2\theta)^{-m/2}$$

Thus we see that $\alpha U\,2n$ is chi-squared distributed with 2n degrees of free-dom. Extensive tables of chi-square values of m degrees of freedom are available which list $\chi^2(\lambda, m)$ for $0 \le \lambda \le 1$ where $\chi^2(\lambda, m)$ is such that

$$\int_{\chi^2(\alpha, m)}^{\infty} h(w)\ dw = \lambda$$

That is $x^2(\lambda, m)$ is the value above which lies $\lambda\%$ of the area under the chi-squared curve.
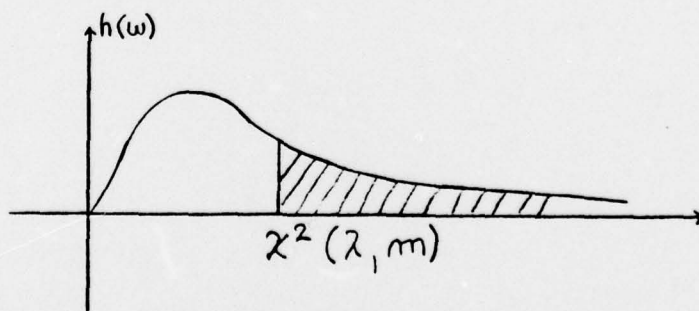
If a desired confidence level λ were set, say $\lambda = .95$, then

$$P[\ \chi^2(.025, 2n) < 2n\alpha U < \chi^2(.975, 2n)\ ] = .95 \quad .$$

After rearrangement we obtain

$$P \left[\frac{\chi^2 (.025, 2n)}{2nU} < \alpha < \frac{\chi^2 (.975, 2n)}{2nU} \right] = 0.95$$

or

$$P \left[\frac{\chi^2 (.025, 2n)}{2 \sum_{i=1}^{n} x_i^\beta} < \alpha < \frac{\chi^2 (.975, 2n)}{2 \sum_{i=1}^{n} x_i^\beta} \right] = 0.95$$

More generally we have for any confidence level λ

$$P \left[\frac{\chi^2 (1 + \frac{\lambda}{2}, 2n)}{2 \sum_{i=1}^{n} x_i^\beta} < \alpha < \frac{\chi^2 (1 - \frac{\lambda}{2}, 2n)}{2 \sum_{i=1}^{n} x_i^\beta} \right] = \lambda .$$

5. The Rayleigh Distribution

A random variable, X, is said to be Rayleigh distributed if its pdf is of
the form

$$f(x) = \begin{cases} \dfrac{2x}{R} \ e^{-x^2/R} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

where $R > 0$.

This type of random variable has been used by many researchers (15). Longuet-
Higgins (8) and Ochi (9) showed that if the normal stochastic process govern-
ing wave displacement from the mean is a narrow-band process (that is one where
$x(t) = A(t) \cos (w_0 t + E(t))$ and E is small). Then the envelope A of the dis-
placements in each small twenty-minute sample of data is Rayleigh distributed
with parameter R changing from sample record to sample record. Since the peak-
to-trough values of such a process may be treated as twice the A values; the
values of the peak-to-trough stress or wave heights per reversal are taken as
Rayleigh distributed random variables with
$R = \dfrac{1}{n} \sum\limits_{i=1}^{n} x_i^2$, the mean square of the sample.

If we compare this $f(x)$ to that of section 3, we quickly see that the
Rayleigh variable is nothing more than the two-parameter Weibull random variable
with $c = 1/R$ and $\beta = 2$. Thus all the Rayleigh variable's properties follow
from before.

$$\mu = \sqrt{R} \ \ \Gamma(1 + \tfrac{1}{2})$$
$$\sigma^2 = R [\ \Gamma (1+1) - \Gamma^2 (1 + \tfrac{1}{2}) \ ] \ .$$

From the appendix $\Gamma (3/2) = \sqrt{\tfrac{\pi}{2}}$ and $\Gamma (2) = 1$ so that

$$\mu = \frac{\sqrt{\pi R}}{2} \ , \quad \sigma^2 = R [\ 1 - \frac{\pi}{4} \ ] \quad .$$

The moment generating function of the Rayleigh variable then is

$$M_X(\theta) = \sum_{n=0}^{\infty} \frac{\theta^n}{n!} \ R^{n/2} \ \ \Gamma (1 + \frac{n}{2}) \quad .$$

This form of the MGF is not too easily dealt with and perhaps a more convenient form is

$$M_X(\theta) = 1 + \frac{\sqrt{R\pi}}{2} \theta \, e^{R\theta^2/4} [1 + erf(\sqrt{R}\,\frac{\theta}{2})]$$

which may be obtained by direct calculation of $E(e^{X\theta})$.

Here $erf(\theta)$ is the error function discussed in the appendix.

The maximum likelihood estimate of R then is

$$R = \frac{1}{n} \sum_{i=1}^{n} x_i^2 ,$$

which is the previously used value. Indeed the maximum likelihood estimate in this case represents, being the sum of squares of displacements, the average energy over the sample period. The λ confidence interval for R then is

$$P\left[\frac{2\sum_{i=1}^{n} x_i^2}{\chi^2(1-\lambda/2,2n)} < R < \frac{2\sum_{i=1}^{n} x_i^2}{\chi^2(1+\lambda/2,\,2n)} \right] = \lambda .$$

This may then be used to develop a confidence for μ and σ^2 through $\mu = \frac{\sqrt{\pi R}}{2}$, $\sigma^2 = R[1 - \pi/4]$ if the sampled population is indeed Rayleigh rather than take the point estimates of $\mu \approx \frac{1}{n} \sum_{i=1}^{n} x_i$ and $\sigma^2 \approx \frac{1}{n} \sum_{i=1}^{n} [x_i - \frac{1}{n}\sum_{i=1}^{n} x_i]^2$.

Also, since

$$\frac{\sigma^2 + \mu^2}{\mu^2} = \frac{\Gamma(1 + \frac{2}{\beta})}{\Gamma^2(1 + \frac{1}{\beta})}$$

for a Weibull random variable, then it is reasonable to expect that for our point estimates of μ and σ^2, we should have

$$\frac{\sigma^2 + \mu^2}{\mu^2} \approx \frac{\Gamma(2)}{\Gamma^2(\beta/2)} = \frac{4}{\pi} .$$

If this is not true, then the Rayleigh distribution can not fit the data well. If it is roughly true, then a chi-square test (see Freund (1)) will tell how accurate the fit is.

6. The Exponential Distribution

Various writers have investigated the application of different distributions to studies of ship hull stress statistics. In some cases they have found, or assumed, that long-term data (several years) roughly fit an exponential distribution (4) (5). Hence, we discuss the exponential distribution.

A random variable, X, is said to be exponentially distributed if its pdf

$$f(x) = \begin{cases} \dfrac{1}{\eta} \, e^{-\frac{x}{\eta}} & x > 0 \\[2mm] 0 & x \leq 0 \end{cases}$$

If we compare this to section 3, we see that the exponential random variable is nothing more than the two-parameter Weibull variable with $\beta = 1$ and $\alpha = 1/\eta$. Hence the mean and variance are

$$\mu = \eta \, \Gamma(2) = \eta$$
$$\sigma^2 = \eta^2 \, [\, \Gamma(3) - \Gamma^2(2) \,] = \eta^2 \quad .$$

The MGF of X has been developed in section 4 as

$$M_X(\theta) = (1 - \eta\theta)^{-1} \quad .$$

The maximum likelihood estimate of η is thus

$$\eta = \frac{1}{n} \sum_{i=1}^{n} x_i \quad ,$$

and the λ confidence interval for η is

$$P \left[\frac{2 \sum_{i=1}^{n} x_i}{\chi^2 (1 - \frac{\lambda}{2}, \, 2n)} < \eta < \frac{2 \sum_{i=1}^{n} x_i}{\chi^2 (1 + \frac{\lambda}{2}, \, 2n)} \right] = \lambda$$

Again, as in section 5, if we use the point estimates of
$\mu \approx \frac{1}{k} [x_1 + x_2 + \dots + x_k]$ and $\sigma^2 \approx \frac{1}{k} [\, \sum_{i=1}^{K} (x_i - \frac{1}{k} \sum_{i=1}^{K} x_i)^2 \,]$ for a sample of size k, it will only be reasonable to assume that an exponential random variable is at work if $\sigma^2/\mu^2 \approx 1$

Application to ship stresses is discussed in Part 2.

7. The Normal Distribution

Although the normal distribution is not directly applicable to ship stress statistics, it will be shown in Section 2 to be applicable indirectly to the problem. Accordingly, in this section we wish to list the properties of the normal random variable and compare it to those of the Weibull random. It has been conjectured that perhaps the two random are indeed one and the same or at least their density curves are so close that they may be treated as the same.

A random variable, X, is said to be normally distributed if it's pdf

$$f(x) = \frac{1}{\sqrt{2\pi}\ \sigma}\ e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \qquad -\infty < x < \infty$$

It is a two-parameter distribution whose parameters are actually the mean and standard deviation. If $z = \frac{x-\mu}{\sigma}$ then the pdf of z is

$$g(z) = \frac{1}{\sqrt{2\pi}}\ e^{-z^2/2} \qquad -\infty < z < \infty$$

which is again a normal distribution which is referred to as standard normal. Although the cumulative distribution function can not be directly calculated, tables of its values for the standard normal case are readily available, and these may quickly be related to any normal random variable.

We first obtain the MGF of the standard normal variable z.

$$M_z(\theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{z\theta}\ e^{-z^2/2}\ dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-1/2\ [z^2 - 2z\theta]}\ dz$$

We complete the square of the exponent and have

$$M_z(\theta) = e^{\theta^2/2}\ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-1/2\ [z^2 - 2z\theta + \theta^2]}\ dz$$

or

$$M_z(\theta) = e^{\theta^2/2}$$

Since the general normal random variable X may be written as $X = \sigma(z + \frac{\mu}{\sigma})$, property 3 again yields

$$M_X(\theta) = e^{\frac{\theta^2\sigma^2}{2} + \mu\theta}$$

The normal distribution is the most frequently used and most important probability distribution because of the central limit theorem which guarantees that the sum of a large number of independent random variables may be treated as if the sum were normally distributed. Property 4 of moment generating functions enables us to apply the central limit theorem to approximate many common random variables via the normal distribution. For example the aforementioned chi-squared distribution with m degrees of freedom had MGF of $(1 - 2\theta)^{-m/2}$. When $m=2n$, this is the product of n MGFs of $(1 - 2\theta)^{-1}$ which is the MGF of an exponential random variable. Thus, the chi-square variable may be viewed as the sum of n independent exponentially distributed random variables and hence, via the central limit theorem, approximated by normal random variable with the same mean and variance.

In order to see how well the normal and Weibull distributions agree, we must first decide which form of the Weibull to use. Since the normal may take on negative values, the three-parameter form initially seems more appropriate. However, since the discrepancy between the two forms applies to energy states which are always positive, the two-parameter form seems to be the one to compare to the normal. Thus we compare the moment generating functions:

$$e^{\frac{\theta^2\sigma^2}{2} + \mu\theta} \qquad \text{(normal)}$$

and

$$\sum_{n=0}^{\infty} \frac{\theta^n}{n!} \alpha^{-n/\beta} \Gamma\left(1 + \frac{n}{\beta}\right) \qquad \text{(two-parameter Weibull)}$$

Even when the normal MGF is expanded as a series, the presence of $\Gamma(1 + \frac{n}{\beta})$ in the Weibull MGF make the two functions sufficiently different that it appears

there is no hope of making them agree. Thus, the normal and two-parameter Weibull distributions are two distinctly different distributions!

Let us still examine the graphs of these two density functions over the range of values usually dealt with. The mean and variances of a probability density function characterize the centering and spread from the center of the probability weight. The higher moments (3rd, 4th... etc.) of the distribution characterize the skewness of the density about the mean. (That is the tendency to flatness in the density curve on one side of the mean.) If two probability density curves are to be roughly the same then it is clear that in addition to their means and variances being equal, as many higher moments as possible must be also equal. Thus, if the normal and Weibull are approximately equal, we must have

$$\mu = \alpha^{-1/\beta} \, \Gamma(1 + \frac{1}{\beta}) \qquad \qquad \text{(1st moment)}$$

$$\sigma^2 + \mu^2 = \alpha^{-2/\beta} \Gamma(1 + \frac{2}{\beta}) \qquad \qquad \text{(2nd moment)}$$

$$3\mu\sigma^2 + \mu^3 = \alpha^{-3/\beta} \, \Gamma(1 + \frac{3}{\beta}) \qquad \qquad \text{(3rd moment)}$$

etc.

When the first two equations are inserted into the third, we get the condition

$$\Gamma(1 + \frac{1}{\beta}) \, [\, 3 \, \Gamma(1 + \frac{2}{\beta}) - 2 \, \Gamma^2(1 + \frac{1}{\beta}) \,] = \Gamma(1 + \frac{3}{\beta})$$

or

$$0 = H(\beta) = \Gamma(1 + \frac{3}{\beta}) - \Gamma(1 + \frac{1}{\beta}) \, [\, 3 \, \Gamma(1 + \frac{2}{\beta}) - 2 \, \Gamma^2(1 + \frac{1}{\beta}) \,]$$

A value of β which makes $H(\beta) \sim 0$, in addition to satisfying the first two equations, will produce a Weibull density quite similar to a normal density.

An investigation of $H(\beta)$ (see Figure 2) shows that $H(\beta) = 0$ for $\beta \sim 3.5$. However, for $3 \leq \beta \leq 10$, $|H(\beta)| < 5 \times 10^{-3}$. Thus, it appears that a Weibull distribution will behave "like" a normal distribution for $3 \leq \beta \leq 10$, and that the closest approximation will be when $\beta \overset{\sim}{\sim} 3.5$. It is reasonable to ask how the 4th and higher moments behave when $H(\beta) = 0$. Unfortunately, when $H(\beta) = 0$,

the 4th moments do not agree. Hence, the Weibull will not tend to a normal exactly, but it can be made to fit up to the 3rd moment. The accompanying graphs (Fig. 3) illustrate how good this fit is.

8. The Voznesensky Distribution

A brief summary of the Voznesensky distribution may be found in Ochi (9). This distribution was developed from an empirical standpoint using the random samples of peak values instead of a spectral analysis.

Voznesensky defined his distribution through its moments. First, he requires a random variable, X, where k^{th} moments were $E(X^k) = \Gamma(1 + \frac{k}{n})$ for some still to be determined n. Next the random variable governing the distribution of peak values per reversal was taken as $Z = X/\Gamma(1 + \frac{1}{n})$. The n could then be determined using point estimates of μ and σ from the equation

$$\frac{\sigma}{\mu} = \frac{\Gamma(1 + 2/n) - \Gamma^2(1 + 1/n)}{\Gamma(1 + 1/n)}$$.

If we compare the moments with those of the Weibull variable in section 3, we see that Voznesensky's X-variable is merely a two-parameter Weibull variable with $\alpha = 1$ and $\beta = n$. His condition for determining n is the condition

$$\frac{\sigma^2 + \mu^2}{\mu^2} = R(\beta)$$

which was discussed before. Hence we shall not further discuss this special case of the Weibull distribution.

9. The Generalized Rayleigh Distribution

The earlier work of Rice (10), is summarized in Ochi (9). Rice examined the stochastic process $x(t)$ governing the displacement from the mean. Assuming $x(t)$ is normal, the distributions and correlations of $\dot{x}(t)$ and $\ddot{x}(t)$ were derived. Let $f(x(t), \dot{x}(t), \ddot{x}(t))$ be the joint density of these three random variables. Maxima of $x(t)$ occur when $\dot{x}(t)$ is zero while $\ddot{x}(t)$ is negative, and probability of peaks of $x(t)$ exceeding level ξ may be obtained by integrating the joint density function with $\dot{x}(t) = 0$, over the range $\xi < x(t)$ and $\ddot{x}(t) < 0$. The chief advantage of this approach is the appearance of the spectral bandwidth parameter ε in the distribution.

When the above process is finished, we arrive at the following density function for X, the maxima of the stochastic process.

$$f(x) = \frac{1}{\sqrt{2\pi}} \; \varepsilon \; e^{-x^2/2\varepsilon^2} + \sqrt{1 - \varepsilon^2} \; xe^{-x^2/2} \int_{-\infty}^{\frac{x\sqrt{1-\varepsilon^2}}{\varepsilon}} e^{-t^2/2} \, dt$$

for
$$-\infty < x < \infty \;, \qquad 0 < \varepsilon < 1$$

A random variable with this density is said to have the generalized Rayleigh distribution. The name follows from the fact that as $\varepsilon \to 0$, and hence the process becomes narrow band, the distribution approaches that of a Rayleigh random variable with R = 2. (Note: In the above density function we have already divided X by its rms value, and hence the additional parameter does not appear in the density function and R is constant). As $\varepsilon \to 1$, however, and the band width widens, the density function approaches the standard normal case. A graph of $f(x)$ for various ε values is in Figure 4.

Ideally, we would like the MGF for the random variable. Although the form of the density is quite involved and initially seems to prohibit a simple integral evaluation, several pages of integral transformations listed in the appendix yield

-19-

$$M_X(\theta) = e^{\theta^2 \epsilon^2/2} + \theta\sqrt{1 - \epsilon^2} \ \sqrt{\pi/2} \ e^{\theta^2/2} \ [\ 1 + \text{erf} \ (\ \frac{\theta\sqrt{1-\epsilon^2}}{\sqrt{2}} \ )\ ].$$

As expected, we see that as $\epsilon \to 0$, the MGF becomes

$$M_X(\theta) = 1 + \theta\sqrt{\pi/2} \ e^{\theta^2/2} \ [\ 1 + \text{erf} \ (\frac{\theta}{\sqrt{2}})\ ]$$

which is that of a Rayleigh random variable with R=2. As $\epsilon \to 1$, the MGF becomes

$$M_X(\theta) = e^{\theta^2/2}$$

which is that of the standard normal random variable.

If the MGF $M_X(\theta)$ is expanded as a powers series in θ, we have

$$M_X(\theta) = 1 + \sqrt{\pi/2} \ \sqrt{1 - \epsilon^2} \ \theta + [2 - \epsilon^2] \ \theta^2/2! + 3\sqrt{\pi/2} \ \sqrt{1 - \epsilon^2} \ \theta^3/3! + \ldots.$$

from which we may easily read off $E(X)$ and $E(X^2)$. Hence,

$$\mu = \sqrt{1 - \epsilon^2} \ \sqrt{\pi/2}$$

$$\sigma^2 = 1 + (1 - \epsilon^2) \ (1 - \pi/2).$$

This of course agrees with the mean and variance of the two limiting cases.

Before we compare this distribution to the Weibull case, we shall make some interesting observations. The MGF $M_X(\theta)$ above, may be written as

$$M_X(\theta) = M_Z \ (\frac{\theta\sqrt{1 - \epsilon^2}}{\sqrt{2}})$$

where Z is a random variable with MGF

$$M_Z(\theta) = e^{\theta^2\epsilon^2/(1-\epsilon^2)} \left\{ 1 + \sqrt{\pi} \ \theta \ e^{-\theta^2} \ [1 + \text{erf}(\theta)] \right\} .$$

Hence, by property 3

$$X = \frac{\sqrt{1-\epsilon^2}}{\sqrt{2}} \ Z.$$

But $M_Z(\theta)$ is the product of

$$M_W(\theta) = e^{\theta^2\epsilon^2/(1-\epsilon^2)}$$

and

$$M_Y(\theta) = 1 + \sqrt{\pi} \ \theta \ e^{\theta^2} \ [1 + \text{erf} \ (\theta)].$$

$M_W(\theta)$ is the MGF of a normal random variable with $\mu = 0$ and $\sigma^2 = 2\epsilon^2/(1-\epsilon^2)$.

Hence W has density

$$h(w) = \frac{\sqrt{1-\epsilon^2}}{\sqrt{2\pi}\sqrt{2}\,\epsilon} \; e^{-\frac{w^2(1-\epsilon^2)}{4\epsilon^2}} \qquad\qquad -\infty < w < \infty$$

$M_Y(\theta)$ is clearly the MGF of a Rayleigh random variable with R=2 and hence density

$$g(y) = \begin{cases} y \; e^{-y^2/2} & y \geq 0 \\ 0 & y < 0 \end{cases}$$

Thus by property 4, the random variable X may be written as the sum of two independent variables U and V

$$X = \sqrt{\frac{1-\epsilon^2}{2}} \; W + \sqrt{\frac{1-\epsilon^2}{2}} \; Y = U + V$$

One variable, U, is normal with $\mu = 0$ and $\sigma^2 = \epsilon^2$. The other, V, is Rayleigh with $\mu = \sqrt{\pi/2}\,\sqrt{1-\epsilon^2}$ and $\sigma^2 = (2 - \pi/2)(1 - \epsilon^2)$. The Chebyshev Theorem states that any random variable has no more than a 1/n chance of falling more than n σ distance from its mean. Thus, for ϵ near zero, the U values have little effect on X, and for ϵ near 1, the V values have little effect. Thus, the more narrow-band the stochastic process becomes, the more the distribution of the maxima becomes Rayleigh.

If we compare the generalized Rayleigh case to the Weibull case, we immediately see that the three-parameter Weibull distribution must be used since X may take on negative values. Since the form of MGFs are again quite distinct as in section 7, there is no hope of showing that the generalized Rayleigh variable is really a three-parameter Weibull variable with γ, b, a related to ϵ. However, since the Weibull distribution does have three parameters, γ, b, a, these may be selected as functions of ϵ to require that the first three moments of the two distributions fit. This will make the mean, variance and skewness coincide again, and hence the two curves would be quite similar to the eye.

Thus we want

$$\sqrt{\pi/2}\ \sqrt{1-\epsilon^2}\ =\ b\ \Gamma(1+\frac{1}{\gamma})\ +\ a\ =\ E(X)$$

$$2-\epsilon^2\ =\ a^2+2ab\ \Gamma(1+\frac{1}{\gamma})+b^2\ \Gamma(1+\frac{2}{\gamma})\ =\ E(X^2)$$

$$3\sqrt{\pi/2}\ \sqrt{1-\epsilon^2}\ =\ b^3\ \Gamma(1+\frac{3}{\gamma})+3\ ab^2\ \Gamma(1+\frac{2}{\gamma})+3\ a^2b\ \Gamma(1+\frac{1}{\gamma})$$

$$+\ a^3\ =\ E(X^3)$$

Now if $\quad b=\sqrt{1-\epsilon^2}\ \beta\ ,\quad a=\sqrt{1-\epsilon^2}\ \alpha$

The equations become

$$\sqrt{\pi/2}\ =\ \beta\ \Gamma(1+\frac{1}{\gamma})\ +\ \alpha$$

$$1+\frac{1}{1-\epsilon^2}\ =\ \alpha^2+2\alpha\beta\ \Gamma(1+\frac{1}{\gamma})+\beta^2\ \Gamma(1+\frac{2}{\gamma})$$

$$3\sqrt{\pi/2}\ =\ \beta^3\ \Gamma(1+\frac{3}{\gamma})+3\ \alpha\ \beta^2\ \Gamma(1+\frac{2}{\gamma})+3\ \alpha^2\ \beta\ \Gamma(1+\frac{1}{\gamma})+\alpha^3$$

These simplify to

i)
$$\frac{[\ \Gamma(1+\frac{2}{\gamma})-\Gamma^2(1+\frac{1}{\gamma})\ ]^{3/2}}{\Gamma(1+\frac{3}{\gamma})-3\Gamma(1+\frac{1}{\gamma})\ \Gamma(1+\frac{2}{\gamma})+2\ \Gamma^3(1+\frac{1}{\gamma})}\ =$$

$$\frac{[\ (1-\frac{\pi}{2})+\frac{1}{1-\epsilon^2}\ ]^{3/2}}{(\pi-3)\ \sqrt{\frac{\pi}{2}}}$$

ii)
$$\beta\ =\ \sqrt{\frac{(1-\frac{\pi}{2})+\frac{1}{1-\epsilon^2}}{\Gamma(1+\frac{2}{\gamma})-\Gamma^2(1+\frac{1}{\gamma})}}$$

iii)
$$\alpha\ =\ \sqrt{\frac{\pi}{2}}\ -\ \beta\ \Gamma(1+\frac{1}{\gamma})$$

Thus a "good fit" of the Weibull density to the generalized Rayleigh distribution may be achieved with the dependence on ϵ, the spectral width, clearly brought out.

10. Further Weibull Properties (two-parameter)

A) The Mode:

The mode of a random variable X is that value which is most likely to occur and will be the most frequently observed value in a large sample. The mode will occur when the pdf $f(x)$, or equivalent, $\ln f(x)$ achieves its maximum. Now

$$\ln f(x) = \ln \alpha \beta + (\beta - 1) \ln x - \alpha x^\beta$$

$$\frac{d \ln f(x)}{dx} = (\beta - 1) \left[\frac{1}{x} - \frac{\alpha \beta}{\beta - 1} x^{\beta - 1} \right]$$

which vanishes when

$$x = \left[\frac{\beta - 1}{\alpha \beta} \right]^{1/\beta}$$

B) The p-th value x_p:

x_p is the value above which lies p% of the probability weight, or above which p% of the values in a large sample will lie. Thus

$$\int_{x_p}^{\infty} f(t)\, dt = p$$

or

$$e^{-\alpha x_p^\beta} = P$$

Hence

$$\alpha x_p^\beta = \ln (1/p)$$

and

$$x_p = \left[\frac{\ln (1/p)}{\alpha} \right]$$

C) Order Statistics:

If a sample of n independent random variables is taken a large number of times, a histogram of the maximum values can be developed. The maximum value of a sample of size n is itself a random variable, X_{max}, with a distribution. Now requiring $X_{max} \leq X$ is equivalent to requiring that each of the n sampled random variables $X_i \leq x$ for $i = 1, 2, \ldots, n$. Thus if $F(x)$ is the edf of the sampled population, then

$$P(X_{max} \le x) = F(x)^n$$

is the edf of X_{max} and

$$n \, F(x)^{n-1} \, f(x)$$

is the pdf.

Thus for the Weibull random variable the density function of X_{max} is

$$\alpha \, \beta \, n \, [\, 1 - e^{-\alpha x^\beta} \,] \, x^{\beta-1} \, e^{-\alpha x^\beta} \qquad\qquad 0 < x$$

This can be maximized easily to find its mode. This will be the most frequently occurring maximum value observed in many repeats of n samples. The probability of a maximum exceeding this value is so large as to be of little design interest.

11. Numerical Solution

Finally, mention should be made of the possibility of obtaining a long-term distribution by numerically summing a large number of short-term distributions, such as the Rayleigh (6). In this case it is unnecessary that the resulting distribution fit any particular function. Of course, the method can be refined by using the generalized Rayleigh distribution for the short-term data, having an additional bandwidth parameter, ϵ, as discussed in section 9 (6).

Available ship stress data are usually in the form of short-term (20 to 30 minutes) records taken automatically ever four hours. If the peak-to-trough stress variations are assumed to be Rayleigh distributed, then the individual records can be characterized by their Rayleigh parameters, R (rms values). Hence, one way to proceed is to determine a suitable distribution function to describe the R, which may be treated as a random variable with pdf g(R). If X is the measured stress or wave height, the Rayleigh distribution obtained from each record may be considered a conditional distribution of X given R which we denote by f(x/R). Then f(x/R)g(R) is the joint density of X and R, and the density of X alone regardless of R may be obtained from:

$$\int_0^\infty f(x/R) \ g(R) \ dR$$

This is what is referred to as the long-term distribution.

Examples will be presented in Part 2.

The long-term distribution of stress or wave heights per reversal has also been obtained as follows (11), (12), (13), (14). First the data are classified according to weather severity. Within each weather group the twenty-minute records each yield a Rayleigh distribution but with different R values. The R values are then taken as a sample of the distribution of ship stresses within that weather group, and long-term distributions are obtained for each weather group. The results for all weather groups are then weighted according to chances of weather occurences and added. Examples are given in Part 2.

-25-

This approach has the advantage of relating the observed data to the physical cause--the sea conditions--rather than relying on the adoption of a particular distribution function that happens to match the data at low N values. This is claimed not only to result in more reliable extrapolation of the data to large values of N but permits comparison of ships on different services by reducing results to the same "standard" or typical weather conditions.

Within a given weather condition, Nordenstrom has claimed (4) the parameter of short term Rayleigh distribution follows a Weibull distribution, while Webb researchers (11), (6) have claimed that the normal distribution is at work. If one examines a sample of the parameter X_1, X_2,...X_n, using the point estimates of μ and σ^2 to estimate α and β discussed above, the β either does or does not fall in the range where the $H(\beta) = 0$ and where the two distributions are close. The $R(\beta)$ in section 3 which determines β must be close to $R(3.5) = 1.1002$, if a normal is also to fit the data fairly well. Since $3 \leq \beta \leq 10$ seems to give fairly normal approximations and (see graph) since in this range $1.132 \leq R(\beta) \leq 1.014$, it appears that a normal and Weibull distribution may be quite similar over a broad range of α.

A few examples run on the weather data in the appendix of SSC-196 seem to indicate that the weather groups III, IV, V each yield a β within this range, but that the data for groups I, II yield β falling slightly outside the broad range. The lumped weather data also does not fall within the β range. Since $R(\beta) = \dfrac{\sigma^2 + \mu^2}{\mu^2}$, the "best" β will occur when $\sigma^2/\mu^2 \sim .1002$, it seems reasonable that the higher weather groups with the larger μ values will yield β's falling within a range where the normal and Weibull densities are close,

12. Comments and Conclusions

The application of the Weibull distribution to problems in ship research appears particularly promising. Under the assumptions that a normal "stochastic" process is initially at work, a reasonable assumption in light of the central limit theorem, the distributions of the variables of interest may very well be special cases of the Weibull distribution. Those which are not directly Weibull may be such that their density curves will be fit quite well by the Weibull density, so that questions based on areas under these curves may be approximately the same. Accordingly, Part 2 of this report will examine actual full-scale statistical data from ship research to determine the applicability of the Weibull distribution, as well as other approaches discussed here.

Part 2

Application of Long-Term Distributions

Part 2
Application of Long-Term Distributions

1.   Introduction

It is the objective of this part of the report to show the application
of the most promising forms of long-term distribution presented in Part 1
to a marine system response to stochastic inputs. Specifically, ship hull
stresses in irregular waves at sea represent a phenomenon of great interest
to ship designers for which considerable full-scale statistical data are
available.

A review of published data revealed that histograms--and/or cumulative
statistics--were available for the following ships:

Esso Malaysia,	190,000-ton deadweight tanker
R.G. Follis,	66,500-ton deadweight tanker
Fotini L,	74,000-ton deadweight bulk carrier
Wolverine State,	15,000-ton general cargo ship

The data describe the total population of cycles of peak-to-trough midship
bending stress obtained over periods of two to three years. They were
recorded automatically for periods of 20-30 minutes every 4 hours.

It was hoped to obtain similar data for the SL-7 container ship
Sealand Maclean, but it was found that the available data were not in
sufficient detail.

In the case of the first ship mentioned above, the tanker Esso
Malaysia computer cards were available at Webb Institute for the entire $1\frac{1}{2}$
years of data collection. (These were obtained from Teledyne Engineering
Services in connection with a project at Webb sponsored by the American
Bureau of Shipping (15).) Hence, it was decided to use these cards to
make a completely independent statistical analysis of the data, rather
than to rely on the previously published results. The results of this
analysis, presented in the next section, were found to agree quite well
with the published figures (15).

It should be noted that for the analysis of ship stresses the longest time for which data have been collected on any one ship is about three years, whereas a typical ship's lifetime is 20 to 25 years. Hence, the real problem is to obtain a cumulative distribution--or probability model--that can be extrapolated by a time factor of 10 or more. Basically, there are two approaches, one to find a long-term distribution that describes all the data (stress reversals) and the other is to assume that all short-term data (20-minute records) can be described by Rayleigh distributions and that the long-term distribution can be obtained by summing up these short-term distributions. In the latter case it is not necessary to obtain a formal long-term distribution function.

In addition to the Esso Malaysia analysis presented in the next section, long-term trends for the other ships listed above will be presented in the following section. In all cases the applicability of the Weibull distribution discussed in Part 1 will be considered.

2. Esso Malaysia Analysis

Stress Data

Bending stress data from 24 instrumented voyages of the supertanker Esso Malaysia were analyzed. Stresses were recorded in 20-minute records ("intervals"), normally one interval per 4-hour watch. The reduced data population for the 24 voyages, after discarding intervals containing unusable data, consisted of 3589 analyzable records. Information recorded in each interval included:

1. Ship, wind and wave direction information.
2. Counts of number of stress reversals (peak-to-peak) falling within each of 16 stress ranges.
3. Total number of stress reversals in the interval.
4. RMS peak-to-peak stress for the interval.
5. Maximum single peak-to-peak stress value in the interval.

Our analysis of these data has been concentrated mostly on the distribution of the stress reversals, item 2 above.

A total of 785,511 stress reversals were counted and grouped into 16 stress ranges within each record by a Probability Analyzer, a small computer used by Teledyne Materials Research to generate the statistical characteristics (digital) of the recorded intervals (analog). Results were recorded on 3589 computer cards (one for each interval), and it was this set of cards that was further processed at Webb Institute for this project.

The stress ranges into which the Probability Analyzer groups the digitized stress reversals could be adjusted to various levels depending on the magnitude of the maximum measured stresses. In the Esso Malaysia analysis, most tape reels had maximum stresses below 8000 PSI, so the analyzer was adjusted to 16 ranges of 500 PSI each. For reels containing higher stresses, ranges of 750 PSI each (maximum stress 12,000 PSI) or 1000 PSI each (maximum stress 16,000 PSI) were used. The flexibility enabled the stress ranges to be chosen to suit the data without placing excessive numbers of reversals in any one range. The grouping of data from the Esso Malaysia is summarized in Table I.

Table I

Summary of Stress Analysis -- Esso Malaysia

Probability Analyzer Max. Setting, PSI	Stress Group Range, PSI	No. Reels	No. Intervals
8000	500	48	2727
12000	750	10	648
16000	1000	3	214
			3589

To get the whole population into one histogram, the numbers of stress reversals falling within each stress range had to be summed. This required regrouping into constant stress ranges. This was done, approximately, by plotting cumulative curves of the three groups of data and adding them, as in Figure 5. The resulting cumulative curve was then read back at stress level ranges of 500 PSI, and the cumulative number of reversals converted into "percent equal to or less than", by dividing number of occurrences by the total number of reversals, N = 785,511. Division by N + 1 instead of N yielded a "plotting position" for the probability paper analysis which makes the first (lowest) and last (highest) measurements symmetrical with respect to the 0% and 100% levels, respectively. This adjustment enabled the most extreme data point to be plotted (4). The resulting cumulative distribution and plotting positions are tabulated in Table II.

Distribution of the Stress Reversals

The probability of exceeding different levels of stress actually measured on the ship can be determined from Table II. It is the "percent greater than" the given stress level, that is:

$$Q(X > X_j) = 1 - P(X \le X_j)$$

where

$Q(X > X_j)$ = probability of exceeding X_j

$P(X \le X_j)$ = probability of not exceeding X_j
$\qquad$ = (cumulative % from Table II) x $\frac{1}{100}$

Table II

Cumulative Stress Reversal Histogram, Esso Malaysia

Stress Level, PSI	Cum. No. Reversals = or <	Cumulative Percent	Prob. Paper Plotting Position Percent	Probability of Exceedance	$-\log_{10} P$
0	0	0	0		
500	306,700	39.04	39.04	0.6096	0.21
1,000	564,600	71.88	71.88	.2812	0.55
1,500	680,740	86.66	86.66	.1334	0.87
2,000	733,900	93.43	93.43	.0657	1.18
2,500	757,600	96.45	96.45	.0355	1.45
3,000	769,780*	98.00	98.00	.0200	1.70
3,500	776,300	98.83	98.83	.0117	1.93
4,000	780,900	99.41	99.41	.0059	2.23
4,500	782,400	99.60	99.60	.0046	2.40
5,000	783,400	99.73	99.73	.0027	2.57
5,500	784,420	99.86	99.86	.0014	2.85
6,000	784,876*	99.919	99.919	.00081	3.09
6,500	785,070	99.944	99.94	.00056	3.25
7,000	785,295	99.973	99.972	.00027	3.57
7,500	785,369	99.982	99.982	.00018	3.74
8,000	785,411	99.987	99.987	.00013	3.89
9,000	785,455*	99.9929	99.9927	.000071	4.15
10,000	785,477	99.9957	99.9955	.000043	4.37
11,000	785,492	99.9976	99.9975	.000024	4.62
12,000	785,506*	99.9994	99.9992	.000006	5.20
13,000	785,510*	99.9999	99.9997	.000001	5.90
13,800**	785,511*	100.	99.9999		

---

*These are exact. Others are close approximations, not exact counts, because of the varying stress level ranges used in the original analysis (see text).

**Highest single value of stress was 13,800 PSI.

The probability of exceedance thus derived from the stress reversal histogram is plotted in Figure 6. For a long-term probability of exceedance of stress levels beyond those actually measured, the curve in Figure 6 must be extrapolated to higher stress levels and smaller probabilities. To do this, a probability distribution model must be determined which fits closely the observed data and therefore gives a reliable estimate of predicted extreme values.

Most attempts at fitting a distribution to the entire population of stress reversals have not produced satisfactory fits at extreme stress levels. However, it appears from Part 1 that one type of extreme value distribution, the Weibull distribution, may satisfactorily describe the entire population of stress reversals. This hypothesis was tested for the Esso Malaysia data by plotting the cumulative stress percentages on Weibull probability paper, on which a true distribution plots as a straight line. Figure 7 shows such a plot, using the stress levels and plotting positions given in Table II. It is clear from this plot that the entire population of stress reversal data is not well fited by a Weibull distribution, since the plotted points describe a curve rather than a straight line. The question remains whether an approximate straight line fit can be determined such that the corresponding Weibull distribution will define a satisfactory long-term curve to extrapolate the measured data. Two visual fits were tried, as shown on the figure. Line 1 was fitted to the entire range of data, and Line 2 favored the high-stress data, where the trend was nearly linear. In both cases the Weibull parameters were estimated from the plotted straight lines. The cumulative Weibull distribution is:

$$Q(x) = e^{-\alpha x^{\beta}} = \text{probability of exceedance, where}$$
$$x = \text{stress level}$$

The Weibull paper is constructed with linear scales of ln x as abscissa (horizontal scale), and $\ln \ln \frac{1}{Q}$ as ordinate (vertical scale). A straight line on the Weibull paper has slope = β and the intercept with the zero axis of ln x is α. That is, the straight line has the equation

$$\ln \ln \frac{1}{Q} = \beta \ln x + \ln \alpha$$

where α and β are the Weibull parameters.

The parameters determined from the Weibull fits are as follows:

Line 1 $\alpha = 0.00060$ $\beta = 1.091$

Line 2 $\alpha = 0.00512$ $\beta = 0.826$

Each of these lines, in addition to the actual curve from the histogram, was used to predict long-term trends. The long-term curves are plotted in Figure 8, which shows, as expected, a poor fit for Line 1. Line 2 is a better fit at high stress levels, but since the number of stress reversals at these levels is so small, the reliability of the Weibull curve as an extrapolator is questionable.

A third Weibull fit was tried, using calculated estimates of α and β determined as explained in Part 1, rather than from a Weibull paper plot. The resulting long-term curve ($\alpha = .00041$, $\beta = 1.192$) was found to be entirely inadequate to represent the data. This numerical technique cannot "bias" the fitted line toward any particular part of the histogram. All three fitted Weibull lines were then subjected to chi-squared goodness of fit tests to the measured data. In all cases the results showed poor to very poor correlation with the data. It is concluded that the Weibull distribution is not an acceptable model of unstratified stress-reversal data.

3. Analysis of Data for Other Ships

The available stress reversal measurements of several other ships were also tested for Weibull fits, and the results are shown in Figure 9. The ships are the R.G. Follis, the Fotini L. and the Wolverine State. It can be seen that the Follis and Fotini stresses exhibit the same characteristic curve as that found for the Esso Malaysia data when plotted on Weibull paper, and the fit to a straight line is just as unsatisfactory. Only the Wolverine State data are well described by a Weibull distribution. Why this should be so is not known, but it is clear that the Weibull distribution cannot be assumed to describe bending stress distribution for all types of ships.

Distribution of Rayleigh Parameters

Since the stress reversals do not seem to be well described by a Weibull distribution, it is of interest to test whether the Rayleigh parameters (rms peak-to-trough or $\sqrt{E}$) of the 20-minute stress records--or the maximum single stress reversals from each record (X_{max})--provide a better fit to a Weibull distribution. Nordenstrom (16) made such an analysis of the Wolverine State data, as shown in Figure 10, where the Weibull distribution of both $\sqrt{E}$ and X_{max} is seen to be good. Since this information was also available from the Esso Malaysia data, plots were made on Weibull paper as shown in Figure 11. As in the case of the stress reversals, these data plot as curves, although the Weibull fit is better for $\sqrt{E}$ than for the stress-reversal data of the Esso Malaysia. Similar rms peak-to-trough stress data have also been determined for the high-speed SL-7 containership Sea-Land McLean, and the distribution plotted in Figure 12, indicating a similar nonlinear trend. (As previously noted, stress reversal data were not available for this ship).

It appears, in summary, that the following observations can be made about the Weibull distribution of ship stress data:

1. Weibull fits to stress reversal data are poor.
2. Weibull fits to $\sqrt{E}$ values are fair to good.
3. Weibull fits to X_{max} values are poor.

Previous conclusions (17) that the Weibull distribution of $\sqrt{E}$ values gave excellent agreement with statistical data were based largely on an analysis of Wolverine State data, which, as we have shown, seems gratuitously to be very well fitted by Weibull distributions. When applied to other vessels, although the agreement is not as good, it appears to be acceptable. Previous work (11) shows that the normal distribution of $\sqrt{E}$ values is also acceptable.

Grouped Data, Normal $\sqrt{E}$ Distribution

The method of extrapolating ship stress data to predict long-term trends, which has been the standard procedure used by researchers at Webb Institute, was described in Part 1 of this report. Briefly repeated, the procedure assumes:

- Normal distribution of $\sqrt{E}$ values within each Beaufort number group, with observed standard deviation for each.
- Rayleigh distribution of peak-to-trough stresses within each record.
- Actual distribution of Beaufort Numbers as experienced in service.

Then three integrations are performed to obtain the long-term curve. This procedure does not require any specific formulation of a distribution function for the long-term curve.

For the four ships for which stress histograms were available, long-term distribution curves computed in this way have been published previously for the Esso Malaysia, R.G. Follis and Fotini L in (15). The results are shown in Figure 13, in which each long-term distribution curve is plotted along with the histogram data points for each ship.

A study of Figure 13 shows that the fit of these "grouped Normal $\sqrt{E}$" long-term curves to the actual measured data is good for the R.G. Follis and excellent for the other two vessels. The agreement between curves and data over the full range of the data establishes confidence in the curve as an extrapolator to lower probabilities. That is, it is reasonable to assume that the probability model used here should correctly describe the stresses over a long period (20 years or more) since it correctly describes the cumulative stress distribution observed over a period of two or three years, provided that the conditions of operation (route, weather distribution, and ship characteristics) remain unchanged.

4. Long-Term Distribution Techniques Compared

The Wolverine State analysis was published in references (17), (11), (13 and (18). A number of different long-term curves were calculated from these reports, based on varying amounts of data available over the years, and on modified instrumentation and analysis procedures described in (18). The long-term curve representing the most comprehensive data population is from (18), reproduced here as Figure 14. Two long-term curves from grouped data, assuming a normal distribution of $\sqrt{E}$ within each weather group, can be seen to fit the histogram data well. The curves are based on 20 and 44 voyages respectively, while the histogram represents 30 voyages. The Weibull distribution of $\sqrt{E}$ (not grouped) shown in Figure 10 is also plotted in Figure 14, and the fit is also acceptable. Finally, the exponential fit to stress reversal data (see Part 1) has also been calculated for the Wolverine State. This distribution plots as a straight line on Figure 14, and since it lies everywhere below the measured data, it is considered unacceptable as an extrapolator.

Of the two methods which give acceptable fits to the histogram data, which is a better or more reliable extrapolator to much longer times? The answer to this question is somewhat speculative, because verification would require a much longer data collection effort to extend the histogram, and the data collected on the Wolverine State represents the most extensive program of its kind. In principle, grouping the stress data by weather severity and analyzing the several groups as independent populations should be more reliable, because the scatter of $\sqrt{E}$ values within any one weather group would be less than that of the entire population taken as a single group. On this basis, the "grouped normal $\sqrt{E}$" technique might be chosen. For numerical analysis, the normal distribution model also has the advantage of having parameters (mean and standard deviation) which are easily calculated from the data without necessitating special plotting, as is necessary in determining the Weibull parameters. Of course, the weather grouping procedure can be combined with the Weibull distribution assumption by describing separate Weibull fits to each group of data, as is done by Nordenstrøm (19). This procedure would have the advantage mentioned above for grouped data, but would still require the special plotting technique.

5. Conclusions

In the first part of this report it has been shown that the Weibull distribution appears to be a particularly useful tool for statistical studies of stochastic processes. Accordingly, characteristics of the distribution were presented for convenient reference, its estimation discussed, and comparisons made with a normal distribution.

It was shown that the Rayleigh, Voznesensky and exponential distributions are special cases of a two-parameter Weibull distribution, with parameters:

Rayleigh	$\alpha = 1/R$	$\beta = 2$
Voznesensky	$\alpha = 1$	$\beta = n$
Exponential	$\alpha = 1/n$	$\beta = 1$

Finally, the generalized Rayleigh distributed random variable is in reality the sum of two independent random variables, one of which is normally distributed while the other is Rayleigh. As the band width varies, the generalized Rayleigh distribution tends to one of these cases. A 3-parameter Weibull distribution may be closely fit to the generalized Rayleigh distribution with the Weibull parameter dependent on the band width.

In Part 2 the application of the Weibull distribution to the problem of ship hull stress statistics led to the following conclusions regarding the goodness of fit:

> Stress reversal data (X) – poor
> Rayleigh parameters (R or $\sqrt{E}$) – acceptable
> Highest stresses per record (X_{max}) – poor

Accordingly, it appears that the numerical solution of the long-term distribution is most appropriate, using either a Weibull or normal distribution of Rayleigh parameters. It is believed to be preferable in extrapolating stress data to long periods of time to classify and analyze stress data by weather groups.

6. Acknowledgments

7. References

1) Freund, J.E., _Mathematical Statistics_, Prentice Hall, 1971.

2) Feller, W., _An Introduction to Probability Theory and Its Applications_, Wiley Press, 1968.

3) Meyer, P.L., _Introduction to Probability and Statistical Applications_, Addison-Wesley, 1972.

4) Nordenstrøm, N., "Methods for Predicting Long Term Distributions of Wave Loads and Probability of Failure for Ships", Det Norske Veritas Report 71-2-S.

5) Mansour, A.E., "Probabilistic Design Concepts in Ship Structural Safety and Reliability", The Society of Naval Architects and Marine Engineers, 1972.

6) Hoffman, D. and Karst, O.J., "The Theory of the Rayleigh Distribution and Some of Its Applications", _Journal of Ship Research_, September 1975.

7) Voznessensky, A.S. and Netsvetaeff, Y.A., "Wave Energy Spectrum for Model Tests", Contribution to 11th ITTC, 1966.

8) Longuet-Higgens, M.S., "On the Statistical Distribution of the Heights of Sea Waves", _Journal of Marine Research_, 1952.

9) Ochi, M.K., "Statistics for Prediction of Ship Performance in a Seaway", _International Shipbuilding Progress_, 1973.

10) Rice, S.O., "Mathematical Analysis of Random Noise", _Bell System Technical Journal_, 1944-45.

11) Band, E.G.U., "Analysis of Ship Data to Predict Long Term Trends of Hull Bending Moments", American Bureau of Shipping, 1966.

12) Bennet, R., Iverson, A., and Nordenstrøm, N., "Results from Full-Scale Measurements and Predictions of Wave Bending Moments Acting on Ships", Report No. 32 of the Swedish Shipbuilding Research Fdn., October 1962.

13) Lewis, E.V., "Predicting Long-Term Distributions of Wave-Induced Bending Moment on Ship Hulls", SNAME Spring Meeting, 1967.

14) Hoffman, D., Williamson, J., and Lewis, E.V., "Correlation of Model and Full-Scale Results in Predicting Wave Bending Moment Trends", SSC Report SSC-233, 1972.

15) Little, R.S., Lewis, E.V., and Bailey, F.C., "A Statistical Study of Wave-Induced Bending Moments on Large Oceangoing Tankers and Bulk Carriers", _Trans._ SNAME, 1971.

16) Nordenstrøm, N., "A Method to Predict Long-Term Distributions of Waves and Wave-Induced Motions and Loads on Ships", DNV Publication No. 81, April 1973.

17) Lewis, E.V., "Long-Term Applications of St. Denis/Pierson Technique to Ship Design", SNAME Seakeeping Symposium, October 1973.

18) Hoffman, D., vanHooff, R. and Lewis, E.V., "Evaluation of Methods for Extrapolation of Ship Bending Stress Data", Ship Structure Committee Report 234, 1972.

19) Nordenstrøm, N., "Further Analysis of Full Scale Measurements of Midship Bending Moments", Chalmers University Report, Gothenburg, May 1965.

20) Abramowitz, M. and Stegun, I.A., Handbook of Mathematical Functions, National Bureau of Standards, Applied Mathematical Series 55, 1964.

21) Bateman, H., Higher Transcendental Function, McGraw-Hill, 1953.

Appendix 1

The Error Function

The error function erf (x) which appeared throughout the main body of this note is defined as

$$erf(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad .$$

The full list of properties of this function may be found along with tables in Abramowitz (20). The most commonly used properties are listed below.

$$\lim_{x \to \infty} erf(x) = 1$$

$$erf(-x) = - erf(x)$$

$$\frac{d\ erf(x)}{dx} = \frac{2}{\sqrt{\pi}} e^{-x^2}$$

The error function may be approximated by the series expansion

$$erf(x) = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{n!\ (2n+1)}$$

The error function is also associated with the complementary error function erfc(x).

$$erfc(x) \equiv 1 - erf(x) = \frac{2}{\sqrt{\pi}} \int_x e^{-t^2} dt$$

we also have

$$\frac{d\ erfc(x)}{dx} = - \frac{2}{\sqrt{\pi}} e^{-x^2}$$

$$erfc(-x) = 1 - erf(-x) = 1 + erf(x)$$

Appendix 2

The Gamma Function

The gamma function $\Gamma(x)$ is defined as

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

This integral converges for $x > 0$, and via integration by parts satisfies the recursive relation

$$\Gamma(x) = (x-1) \ \Gamma(x-1) \quad .$$

Since $\Gamma(1) = \int_0^\infty e^{-t} dt = 1$, for any integer m

$$\Gamma(m) = (m-1)!$$

Also

$$\Gamma(1/2) = \int_0^\infty t^{-1/2} e^{-t} dt$$

$$= 2 \int_0^\infty e^{-u^2} dU \qquad (t = u^2)$$

$$= \sqrt{\pi} \ \operatorname{erf}(\infty) = \sqrt{\pi} \quad .$$

Additional forms of the gamma function are

$$\Gamma(x) = \lim_{n \to \infty} \frac{n! \ n^x}{x(x+1)\dots(x+n)} \qquad \text{for} \quad x \neq 0, -1, -2, \dots$$

and

$$\frac{1}{\Gamma(x)} = x \ e^{\gamma x} \prod_{n=1}^\infty \left[\left(1 + \frac{x}{n}\right) e^{-x/n} \right]$$

where

$$\gamma = .5772156649\dots \quad \text{is Euler's constant.}$$

Although the form of the gamma function is rather involved, extensive tables are available, (Abramowitz (20)). Quick estimations of $\Gamma(x)$ may, however, be made through the recursive property and the minimax polynomial approximation.

$$\Gamma(1 + x) = 1 - .577191652 \ x$$
$$+ .988205891 \ x^2$$

$$- .897056937 \ x^3$$

$$+ .918206857 \ x^4$$

$$- .756704078 \ x^5$$

$$+ .482199394 \ x^6$$

$$- .193527818 \ x^7$$

$$+ .035868343 \ x^8$$

which represents $\Gamma(1 + x)$ for $0 \leq x \leq 1$ within an error of 3×10^{-7}.

Extensive additional properties of the gamma function may be found in Abramowitz (20), Bateman (21).

Appendix 3

Calculation of Moment Generating Function for the
Generalized Rayleigh Distribution

A random variable X is said to have the generalized Rayleigh distribution if its density function is

$$f(x) = \frac{1}{\sqrt{2\pi}} \left[\varepsilon\, e^{-x^2/2\varepsilon^2} + \sqrt{1-\varepsilon^2}\, x\, e^{-x^2/2} \int_{-\infty}^{\frac{x\sqrt{1-\varepsilon^2}}{\varepsilon}} e^{-t^2/2}\, dt \right]$$

$$-\infty < x < \infty$$

The change of variables $x = \frac{\varepsilon}{\sqrt{1-\varepsilon^2}}\, u$ will simplify the calculations, and if $M_u(\theta)$ is the moment generating function of u, then $M_x(\theta)$, the moment generating function of x, will be

$$M_x(\theta) = M_{\frac{\varepsilon u}{\sqrt{1-\varepsilon^2}}}(\theta) = M_u\left(\frac{\varepsilon\theta}{\sqrt{1-\varepsilon^2}}\right)$$

The probability density function of u, g(u), is $f(x(u))\, \frac{\varepsilon}{\sqrt{1-\varepsilon^2}}$

or

$$g(u) = \varepsilon^2 \left\{ \frac{e^{-u^2/2(1-\varepsilon^2)}}{\sqrt{2\pi}\,\sqrt{1-\varepsilon^2}} + \frac{u e^{-u^2/2[\varepsilon^2/1-\varepsilon^2]}}{\sqrt{2\pi}\,\sqrt{1-\varepsilon^2}} \int_{-\infty}^{u} e^{-t^2/2}\, dt \right\}$$

$$M_u(\theta) = \int_{-\infty}^{\infty} e^{\theta u}\, g(u)\, du$$

$$= \varepsilon^2 \left[G(\theta) + \frac{1}{\sqrt{2\pi}\,\sqrt{1-\varepsilon^2}}\, H(\theta) \right]$$

where

$$G(\theta) = \int_{-\infty}^{\infty} \frac{e^{u\theta}\, e^{-u^2/2(1-\varepsilon^2)}}{\sqrt{2\pi}\,\sqrt{1-\varepsilon^2}}\, du$$

and

$$H(\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{u} e^{u\theta}\, u e^{-\frac{u^2}{2}\left[\frac{\varepsilon^2}{1-\varepsilon^2}\right]}\, e^{-\frac{t^2}{2}}\, dt\, du$$

The integral in $G(\theta)$ is quickly recognized as the integral defining the moment generating function of a normally distributed random variable with

$$\mu = 0 \quad \text{and} \quad \sigma^2 = 1 - \varepsilon^2$$

Thus

$$G(\theta) = e^{\frac{\theta^2(1-\varepsilon^2)}{2}}$$

The integral for $H(\theta)$ is more involved. If the order of integration is reversed

$$H(\theta) = \int_{-\infty}^{\infty} \int_{t}^{\infty} ue^{\theta u} \; e^{-\frac{u^2}{2} \left[\frac{\epsilon^2}{1-\epsilon^2} \right]} \; e^{-\frac{t^2}{2}} \; du \; dt$$

which is of the form

$$I = \int_{\infty}^{\infty} \int_{x}^{\infty} ye^{\theta y} \; e^{-1/2 \left[\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} \right]} \; dy \; dx$$

where $y = u$, $x = t$, $\sigma_1^2 = 1$, $\sigma_2^2 = \frac{1-\epsilon^2}{\epsilon^2}$

By completing the square in the integrand we obtain

$$I = e^{\delta^2 \sigma_2^2 / 2} \int_{-\infty}^{\infty} \int_{x}^{\infty} ye^{-x^2 / 2\sigma_1^2} \; e^{-1/2 \left[\frac{y - \theta\sigma_2^2}{\sigma_2^2} \right]^2} \; dy \; dx$$

or

$$I = e^{\theta^2 \sigma_2^2 / 2} \int_{-\infty}^{\infty} \int_{x-\theta\sigma_2^2}^{\infty} [y + \theta\sigma_2^2] \; e^{-1/2 \left[\frac{x^2}{\sigma_2^2} + \frac{y^2}{\sigma_2^2} \right]} \; dy \; dx$$

where y has replaced $y - \theta\sigma_2^2$

Thus $I = I_1 + I_2$ where

$$I_1 = e^{\theta^2 \sigma_2^2 / 2} \int_{-\infty}^{\infty} \int_{x-\theta\sigma_2^2}^{\infty} \theta \; \sigma_2^2 \; e^{-1/2 \left[\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} \right]} \; dy \; dx$$

and

$$I_2 = e^{\theta^2 \sigma_2^2 / 2} \int_{-\infty}^{\infty} \int_{x-\theta\sigma_2^2}^{\infty} y \; e^{-1/2 \left[\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} \right]} \; dy \; dx$$

We now examine I_1 and I_2 and evaluate them through several changes of variables. Under a change of variables $x = f(t,s)$, $y = g(t,s)$ the integral

$$\int_{R} \int F(x,y) \; dx \; dy = \int_{R^*} \int F(f(t,s), g(t,s)) \left| \frac{\partial(x,y)}{\partial(t,s)} \right| dt \; ds$$

whenever $\left| \frac{\partial(x,y)}{\partial(t,s)} \right| \neq 0$

where R* is R expressed in terms of the new variables and

$$\frac{\partial(x,y)}{\partial(t,s)} = \begin{vmatrix} \frac{\partial x}{\partial y} & \frac{\partial x}{\partial s} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial s} \end{vmatrix}$$

First let $\quad t = \dfrac{x}{\sigma_1}, \quad s = \dfrac{y}{\sigma_2} \cdot \left| \dfrac{\partial(x,y)}{\partial(s,t)} \right| = \sigma_1 \sigma_2$

and I_1 becomes

$$I_1 = e^{\theta^2 \sigma_2^2 / 2} \; \theta \, \sigma_1 \sigma_2{}^3 \int\limits_{\infty}^{\infty} \int\limits_{\frac{\sigma_1 t}{\sigma_2} - \sigma_2}^{\infty} e^{-t^2/2} \, e^{-s^2/2} \, ds \, dt$$

We now relate the t,s coordinate system so as to make the line

$s = \dfrac{\sigma_1 t}{\sigma_2} - \theta \sigma_2 \quad$ parallel to the horizontal axis. Thus

$$t = t' \cos \theta - s' \sin \theta$$

$$s = t' \sin \theta - s' \cos \theta$$

where $\quad \tan \theta = \dfrac{\sigma_1}{\sigma_2}, \quad \sin \theta = \dfrac{1}{\sqrt{\sigma_1{}^2 + \sigma_2{}^2}}, \quad \cos \theta = \dfrac{\sigma_2}{\sqrt{\sigma_1{}^2 + \sigma_2{}^2}}$

Hence

$$t = \dfrac{1}{\sqrt{\sigma_1{}^2 + \sigma_2{}^2}} [\sigma_2 t' - \sigma_1 s'], \quad s = \dfrac{1}{\sqrt{\sigma_1{}^2 + \sigma_2{}^2}} [\sigma_1 t' + \sigma_2 s'],$$

$$\dfrac{\partial(s,t)}{\partial(s',t')} = 1 \quad \text{and} \quad e^{-\frac{t^2 + s^2}{2}} = e^{-\frac{(t')^2 + (s')^2}{2}}$$

If we substitute the primed variables into I_1 and the rewrite the integral with the primed notation eliminated we have

$$I_1 = e^{\theta^2 \sigma_2^2 / 2} \; \theta \sigma_1 \sigma_2{}^3 \left[\int_{-\infty}^{\infty} e^{-t^2/2} \, dt \right] \left[\int_{-\frac{\theta \sigma_2^2}{\sqrt{\sigma_1{}^2 + \sigma_2{}^2}}}^{\infty} e^{-s^2/2} \, ds \right]$$

But $\quad \displaystyle\int_{-\infty}^{\infty} e^{-t^2/2} \, dt = \sqrt{2\pi}$

and $\quad \displaystyle\int_{-\frac{\theta \sigma_2^2}{\sqrt{\sigma_1{}^2 + \sigma_2{}^2}}}^{\infty} e^{-s^2/2} \, ds$

$$= \sqrt{2} \int_{-\frac{\theta \sigma_2^2}{\sqrt{2}\sqrt{\sigma_1^2+\sigma_2^2}}}^{\infty} e^{-u^2} \, du$$

$$= \sqrt{2} \int_{-\infty}^{0} e^{-u^2} \, du + \sqrt{2} \int_{0}^{\frac{\theta \sigma_2^2}{\sqrt{2}\sqrt{\sigma_1^2+\sigma_2^2}}} e^{-u^2} \, du$$

$$= \frac{\sqrt{2\pi}}{2} + \sqrt{2} \, \frac{\sqrt{\pi}}{2} \, \operatorname{erf}\left(\frac{\theta \sigma_2^2}{\sqrt{2}\sqrt{\sigma_2^2+\sigma_2^2}}\right)$$

$$= \sqrt{\frac{\pi}{2}} \left[1 + \operatorname{erf}\left(\frac{\theta \sigma_2^2}{\sqrt{2}\sqrt{\sigma_1^2+\sigma_2^2}}\right) \right]$$

Hence

$$I_1 = e^{\theta^2 \sigma_2^2 / 2} \, \theta \, \sigma_1 \, \sigma_2^3 \, \pi \left[1 + \operatorname{erf}\left(\frac{\theta \sigma_2^2}{\sqrt{2}\sqrt{\sigma_1^2+\sigma_2^2}}\right) \right]$$

Next we consider I_2 under the same set of variable changes.

$$I_2 = e^{\theta^2 \sigma_2^2 / 2} \int_{-\infty}^{\infty} \int_{x-\theta\sigma_2^2}^{\infty} y \, e^{-1/2 \left[\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} \right]} \, dy \, dx$$

$$= e^{\theta^2 \sigma_2^2 / 2} \, \sigma \, \sigma^2 \int_{-\infty}^{\infty} \int_{\frac{\sigma_1 t}{\sigma_2} - \theta\sigma_2}^{\infty} s \, e^{-\left[\frac{t^2+s^2}{2}\right]} \, ds \, dt$$

$$= \frac{\sigma_1 \, \sigma_2^2 \, e^{\theta^2 \sigma_2^2 / 2}}{\sqrt{\sigma_1^2+\sigma_2^2}} \left[I_{21} + I_{22} \right]$$

where

$$I_{21} = \sigma_1 \int_{\infty}^{\infty} \int_{-\frac{\theta \sigma_2^2}{\sqrt{\sigma_1^2+\sigma_2^2}}}^{\infty} t \, e^{-1/2 \, [t^2+s^2]} \, ds \, dt$$

$$= \sigma_1 \left\{ \int_{-\infty}^{\infty} t \, e^{-t^2/2} \, dt \right\} \left\{ \int_{\frac{\theta \sigma_2^2}{\sqrt{\sigma_1^2+\sigma_2^2}}}^{\infty} e^{-1/2 \, s^2} \, ds \right\} .$$

$$= 0$$

since

$$\int_{-\infty}^{\infty} t \, e^{-t^2/2} \, dt = 0 \quad,$$

while

$$I_{22} = \sigma_2 \int_{-\infty}^{\infty} \int^{\infty} s \, e^{-1/2\,[t^2 + s^2]} \, ds \, dt$$
$$-\frac{\theta \, \sigma_2^2}{\sqrt{\sigma_2^2 + \sigma_2^2}}$$

But

$$I_{22} = \sigma_2 \left\{ \int_{-\infty}^{\infty} e^{-t^2/2} \right\} \left\{ \int^{\infty} s \, e^{-s^2/2} \, ds \right\}$$
$$-\frac{\theta \, \sigma_2^2}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

$$= \sqrt{2\pi} \quad \sigma_2 \quad e^{-\frac{\theta^2 \, \sigma_2^4}{2(\sigma_1^2 + \sigma_2^2)}}$$

Thus

$$I_2 = \frac{e^{\theta^2 \sigma_2^2 / 2}}{\sqrt{\sigma_1^2 + \sigma_2^2}} \, \sigma_1 \sigma_2^2 \left\{ 0 + \sqrt{2\pi} \, \sigma_2 \, e^{\frac{\theta^2 \, \sigma_2^4}{2(\sigma_2^2 + \sigma_1^2)}} \right\}$$

Finally

$$I = e^{\theta^2 \sigma_2^2 / 2} \, \theta \, \sigma_1 \, \sigma_2^3 \, \pi \left\{ 1 + \mathrm{erf}\left(\frac{\theta \, \sigma_2^2}{2\sqrt{\sigma_1^2 + \sigma_2^2}} \right) \right\}$$

$$+ \frac{\sigma_1 \, \sigma_2^3}{\sqrt{\sigma_1^2 + \sigma_2^2}} \, \sqrt{2\pi} \, e^{\theta^2 \sigma_2^2 / 2} \, \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

If

$$\sigma_1^2 = 1 \, , \quad \sigma_2^2 = \frac{1-\epsilon^2}{\epsilon^2} \, , \quad \text{then} \quad \sigma_1^2 + \sigma_2^2 = 1/\epsilon^2$$

and

$$H(\theta) = e^{\theta^2/2 \, \left[\frac{1-\epsilon^2}{\epsilon^2} \right]} \, \frac{\theta(1-\epsilon^2)^{3/2}}{\epsilon^3} \, \pi \, \left[1 + \mathrm{erf}\left(\frac{\theta}{\sqrt{2}} \, \frac{1-\epsilon^2}{\epsilon} \right) \right]$$

$$+ \frac{(1-\epsilon^2)^{3/2}}{\epsilon^2} \, \sqrt{2\pi} \, e^{\theta^2[1-\epsilon^2]/2}$$

A-8

Thus the moment generating function of U is

$$M_u(\theta) = \varepsilon^2 \left\{ e^{\theta^2(1-\varepsilon^2)\big/ 2} \right\}$$

$$+ \ \theta \ \sqrt{\pi/2} \ \ \left(\frac{1-\varepsilon^2}{\varepsilon^3}\right) \ e^{\frac{\theta^2[1-\varepsilon^2]}{2\varepsilon^2}} \ \left[1 + \text{erf} \ \left(\frac{\theta}{\sqrt{2}} \ \frac{1-\varepsilon^2}{\varepsilon} \right) \right]$$

$$+ \ \frac{1-\varepsilon^2}{\varepsilon^2} \ \ e^{\theta^2[1-\varepsilon^2]\big/ 2}$$

and the moment generating function for $\underline{X}$ is

$$M_x(\theta) = \varepsilon^2 \left\{ e^{\theta^2 \varepsilon^2/2} + \theta \ \sqrt{\frac{\pi}{2}} \ \frac{\sqrt{\varepsilon^2}}{\varepsilon^2} \ e^{\theta^2/2} \ \left[1 + \text{erf} \ \left(\frac{\theta\sqrt{1-\varepsilon^2}}{\sqrt{2}} \right) \right] \right.$$

$$\left. + \ \frac{1-\varepsilon^2}{\varepsilon^2} \ \ e^{\frac{\theta^2 \varepsilon^2}{2}} \right\}$$

$$= \ e^{\frac{\theta^2 \varepsilon^2}{2}} + \theta \ \sqrt{1-\varepsilon^2} \ \sqrt{\pi/2} \ e^{\theta^2/2} \ \left[1 + \text{erf} \ \left(\frac{\theta}{\sqrt{2}} \ \sqrt{1-\varepsilon^2} \right) \right] \ .$$

R(β)
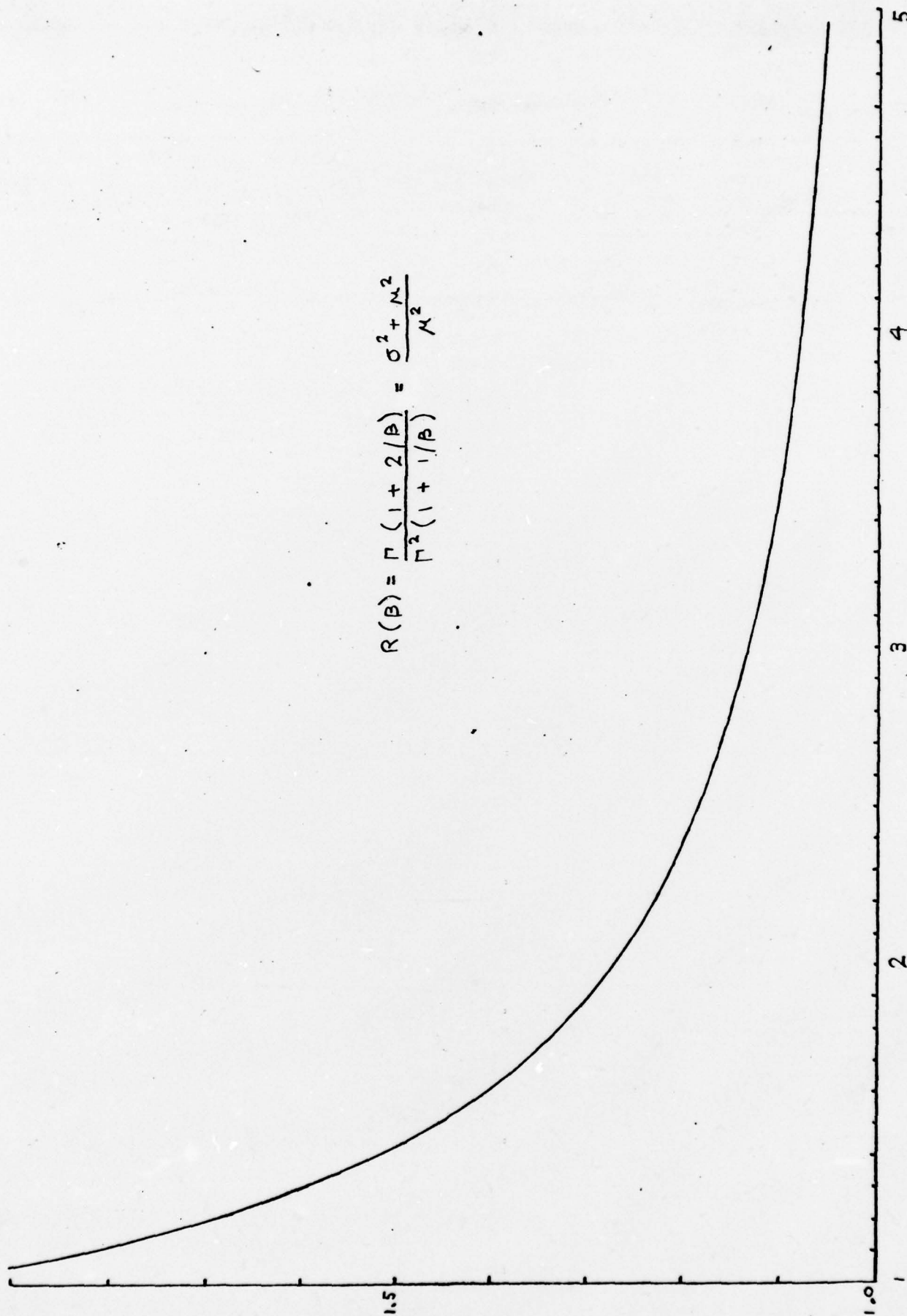
$$R(\beta) = \frac{\Gamma\left(1 + 2/\beta\right)}{\Gamma^2\left(1 + 1/\beta\right)} = \frac{\sigma^2 + \mu^2}{\mu^2}$$

FIGURE 1 R(β) vs. β

$$H(\beta) = \Gamma(1+3/\beta) - \Gamma(1+1/\beta)\left[3\Gamma(1+2/\beta) - 2\Gamma^2(1+1/\beta)\right]$$

FIGURE 2 H(β) vs. β
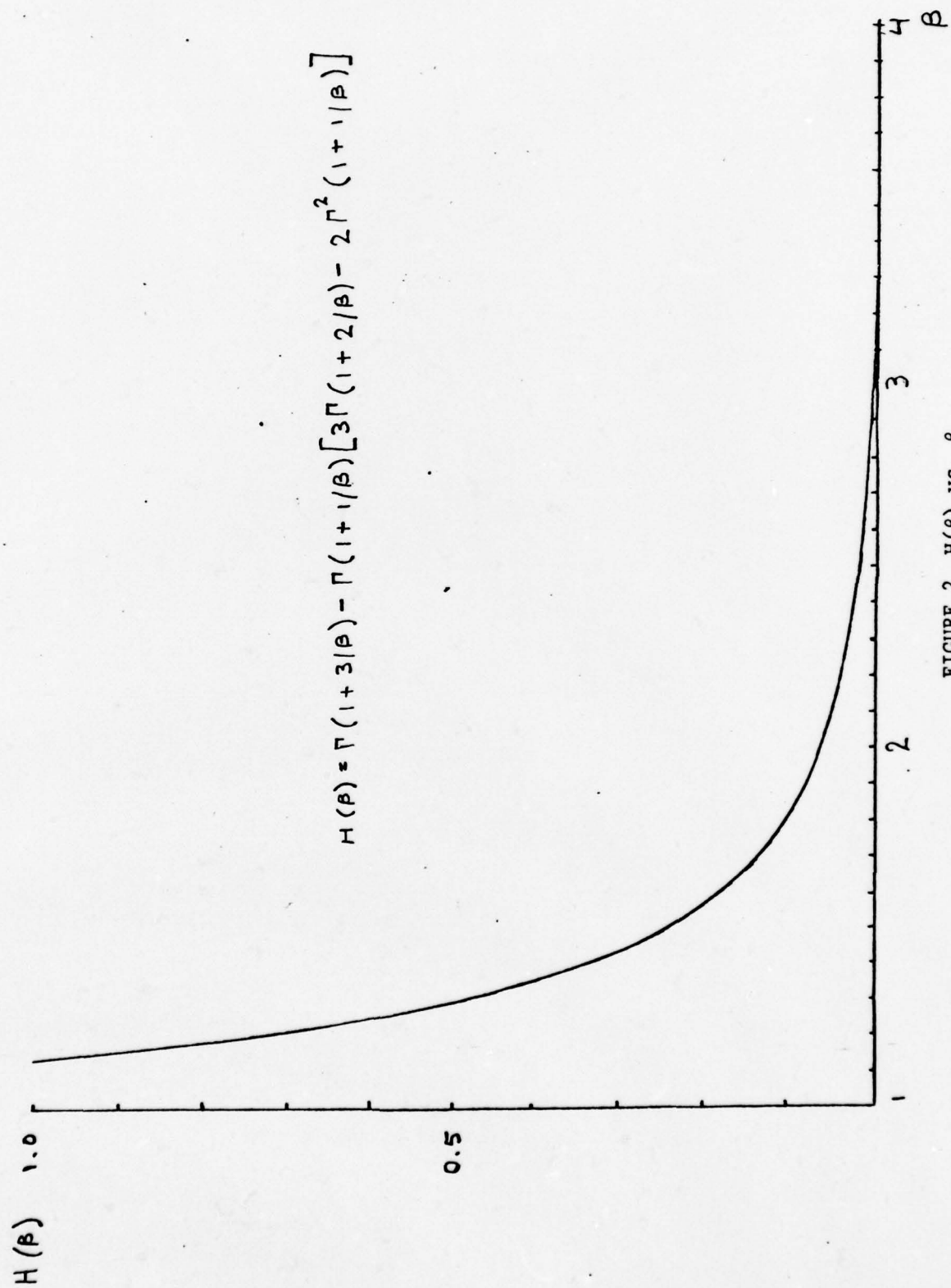
NOTE: Figs. 3-a through 3-d are plots of
normal density curves with the same
μ and σ^2 values. The μ and σ^2 were
calculated for bending stress data
of the S.S. Wolverine State in
various weather groups.
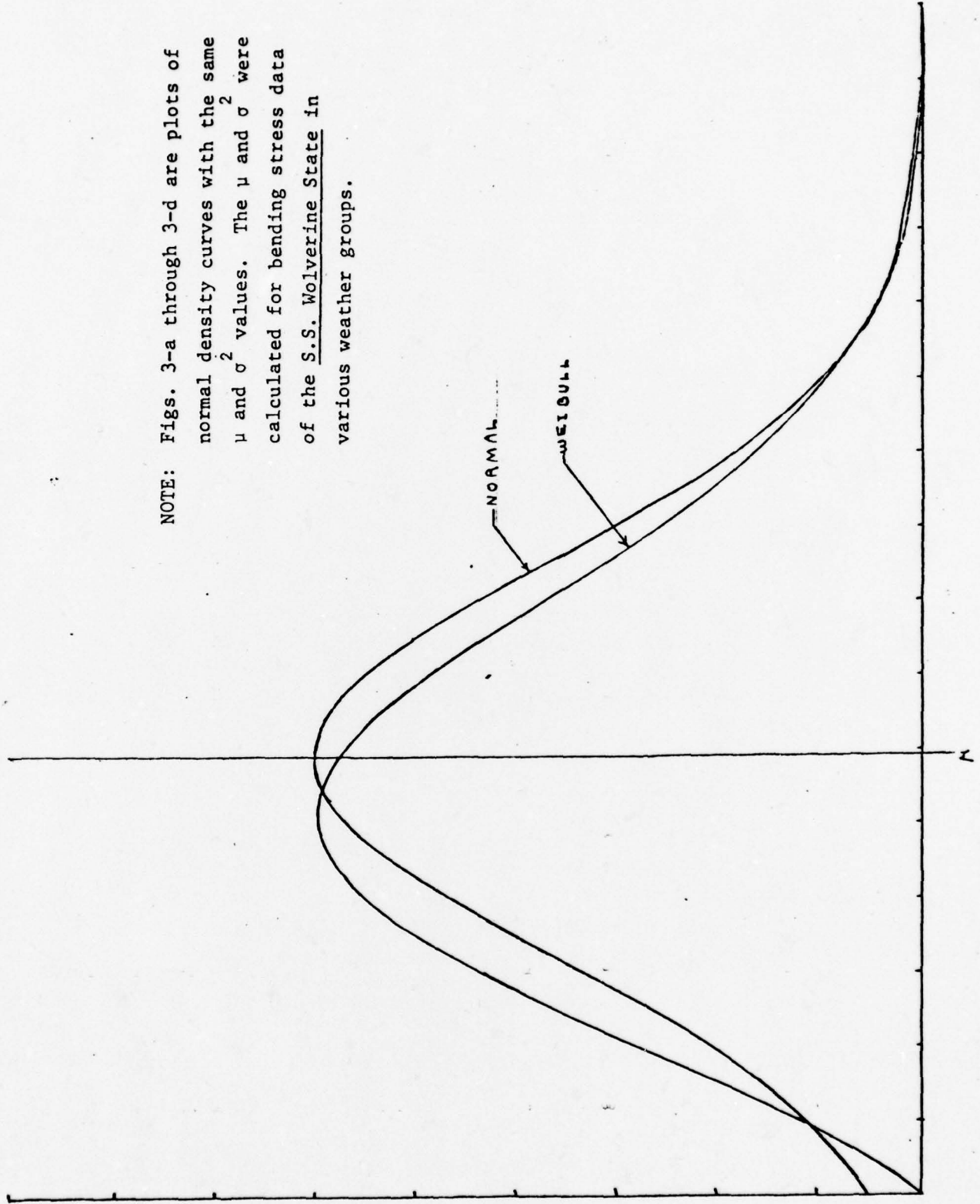
NORMAL

WEIBULL

Fig. 3-a: Normal and Weibull Distributions of Ship Bending Stress Data, Weather Group II.

NORMAL

WEIBULL

μ

Fig. 3-b: Normal and Weibull Distributions of Ship Bending Stress Data, Weather Group III.
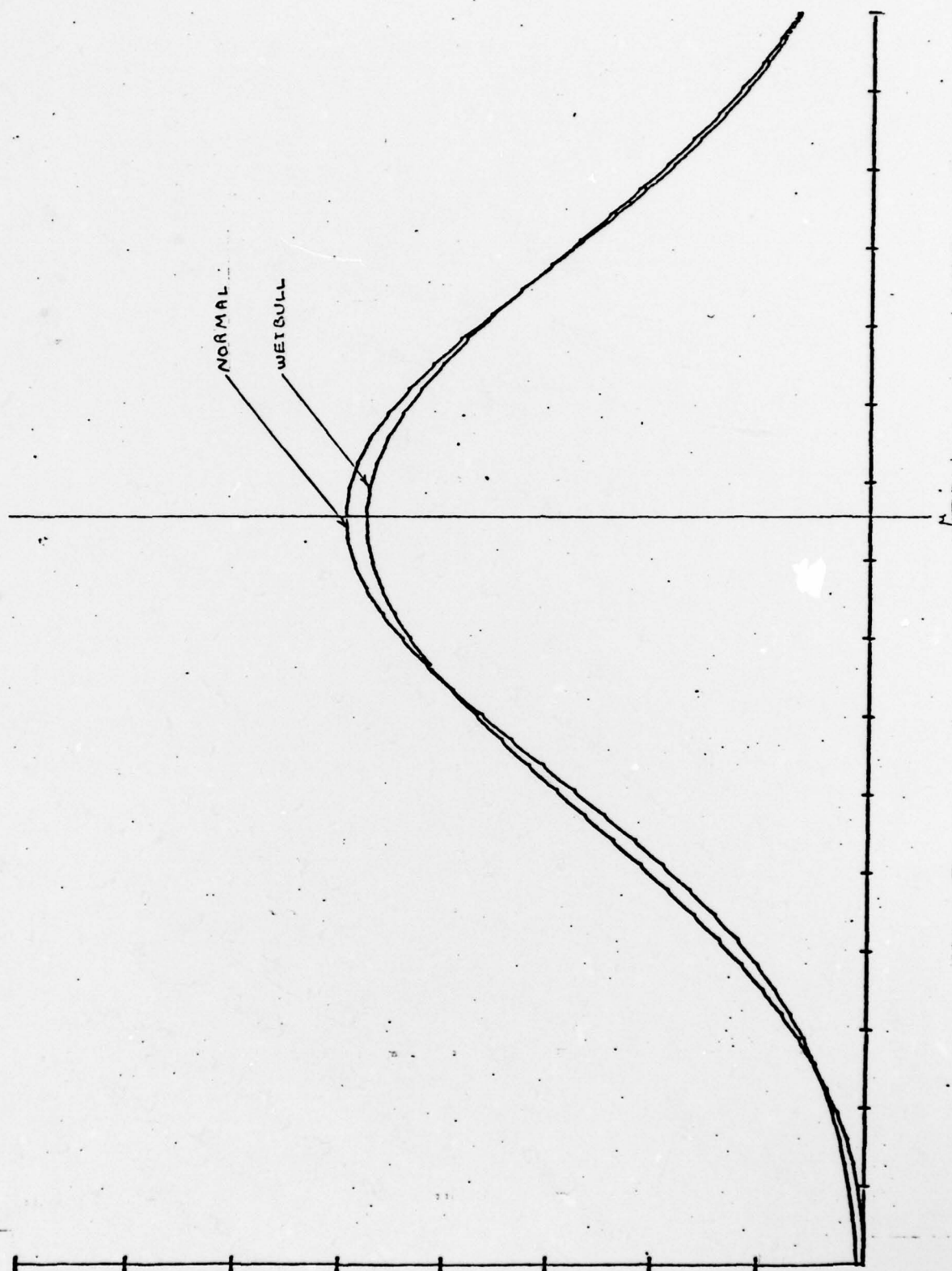
Fig. 3-c: Normal and Weibull Distributions of Ship Bending Stress Data, Weather Group IV.
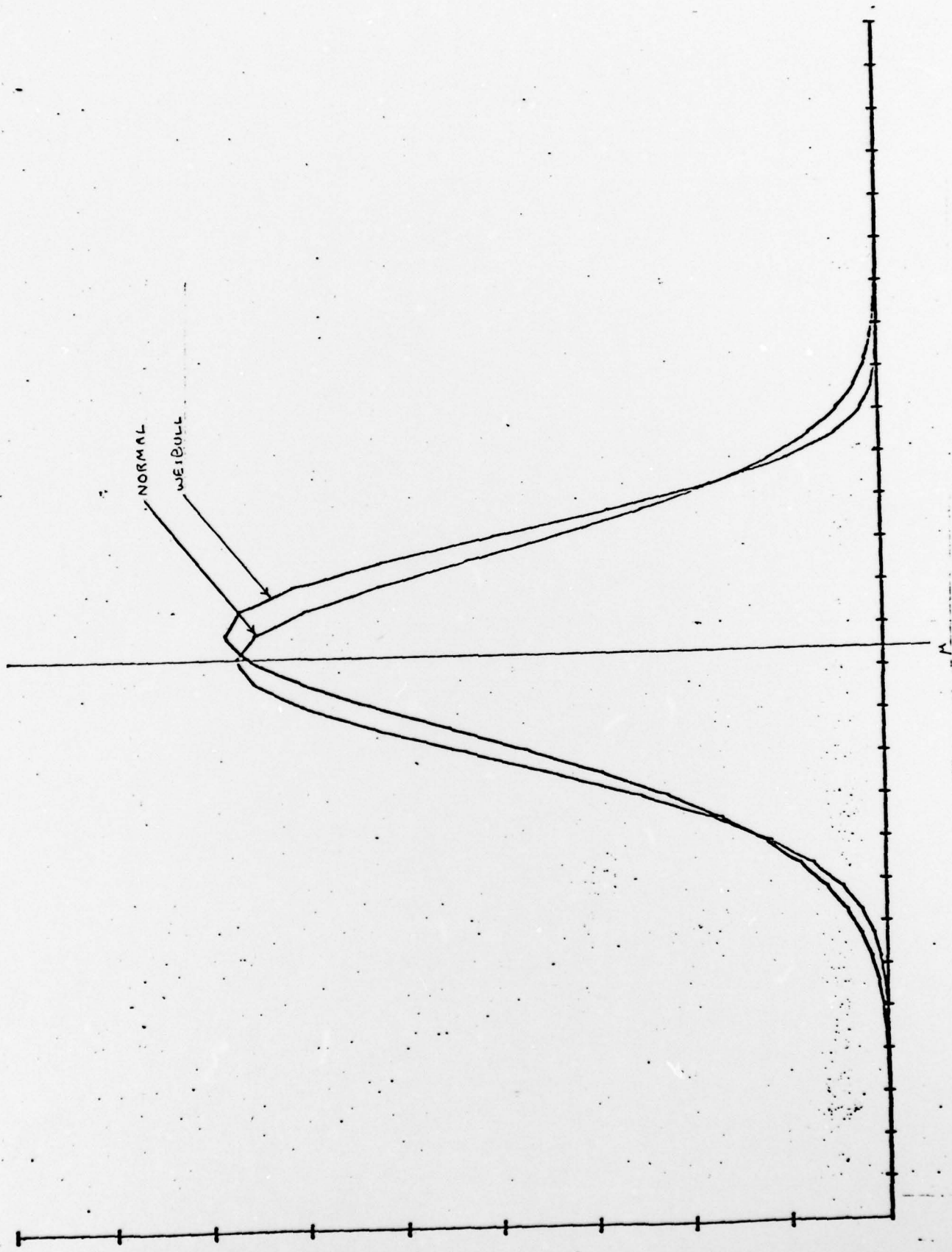
Fig. 3-d: Normal and Weibull Distributions of Ship Bending Stress Data, Weather Group V.
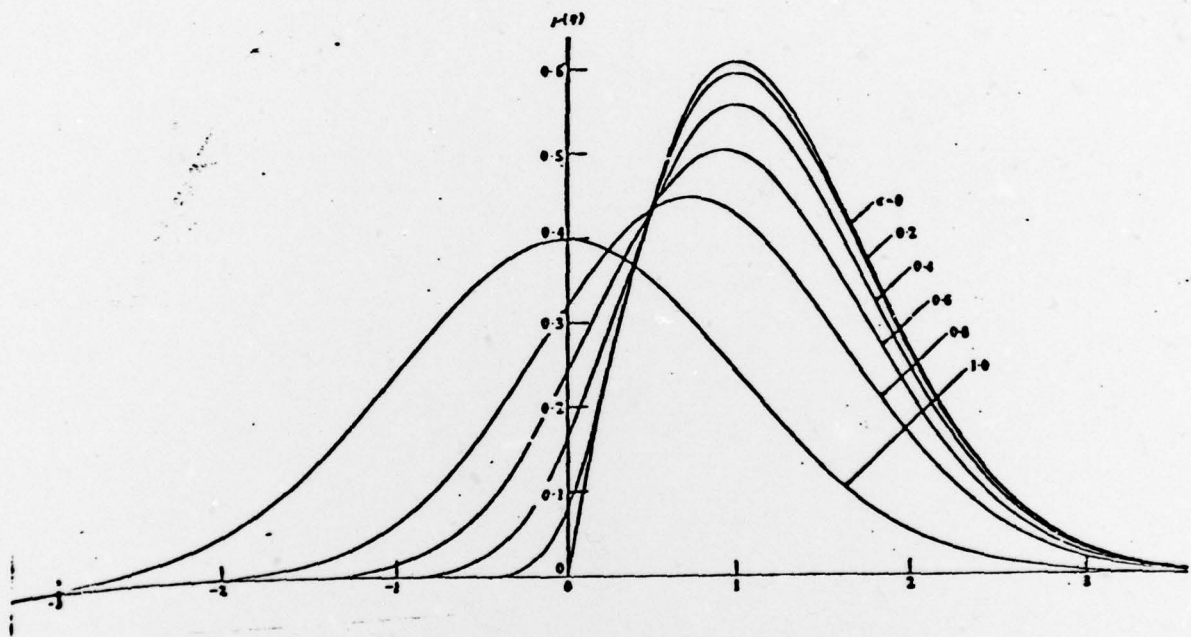
FIGURE 4   The Generalized Rayleigh Distribution
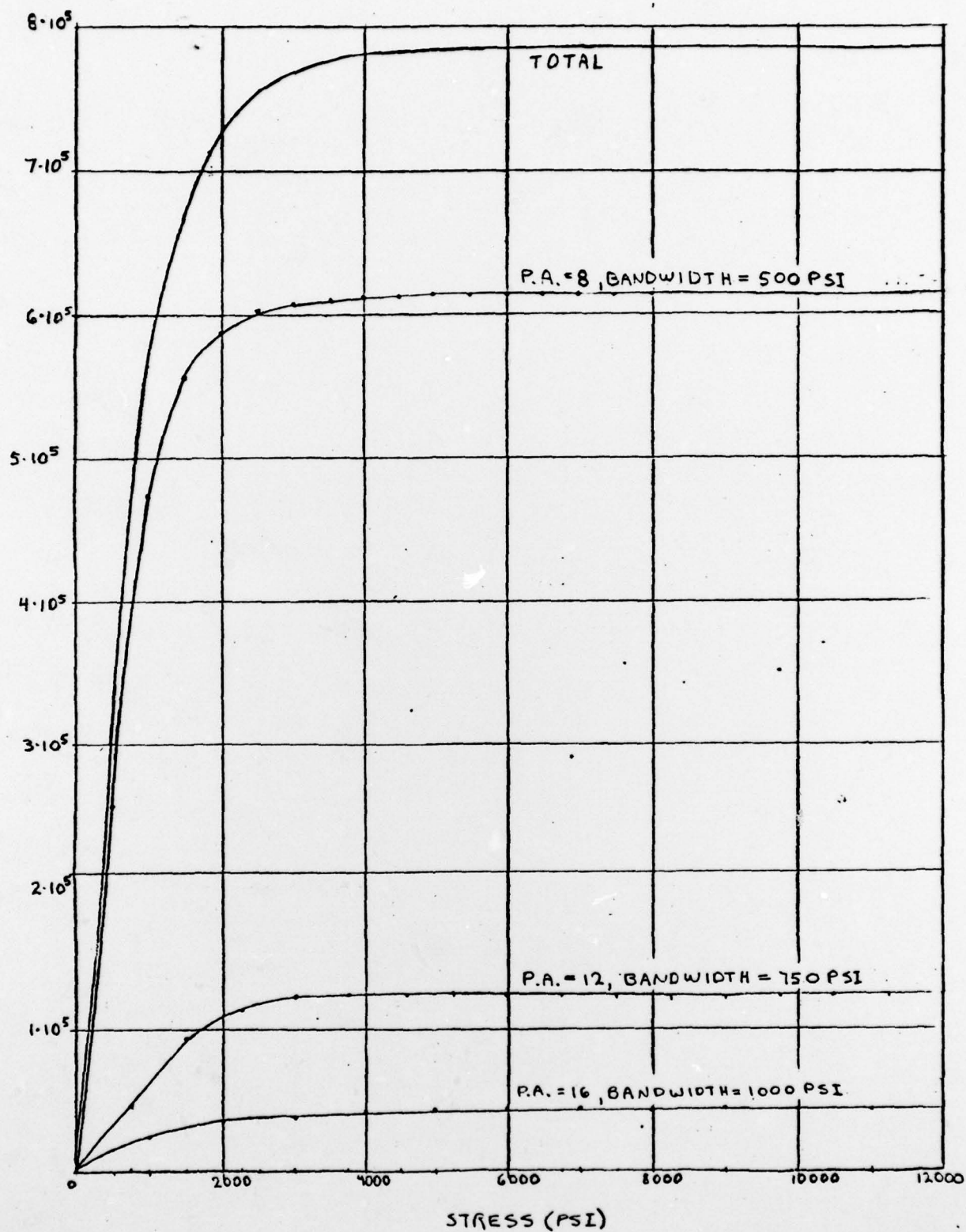
Fig. 5 : Cumulative Histogram of Stress Reversals, Esso Malaysia
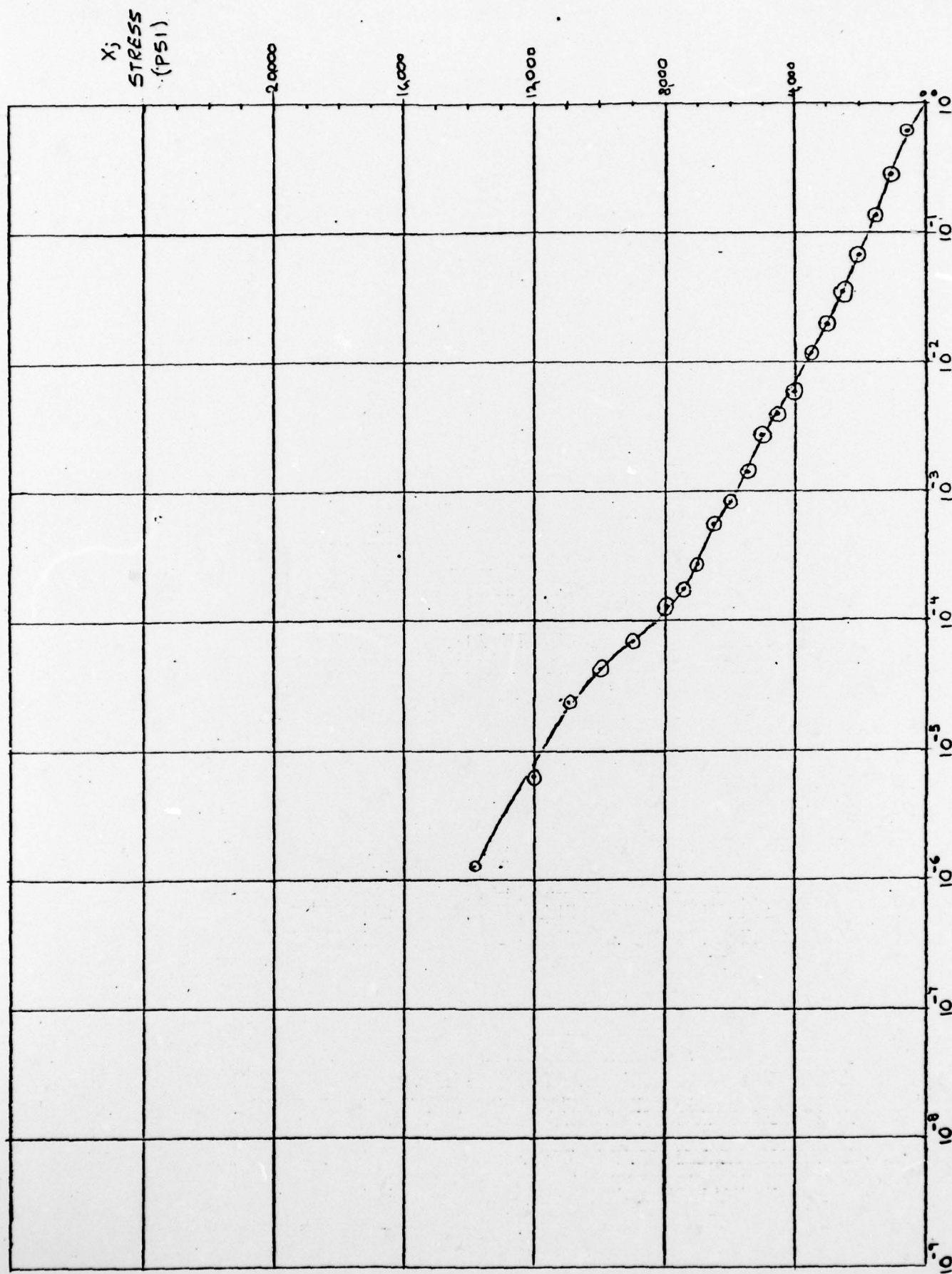
Fig. 6 : Long-Term Distribution of Bending Stress from Histogram, Esso Malaysia
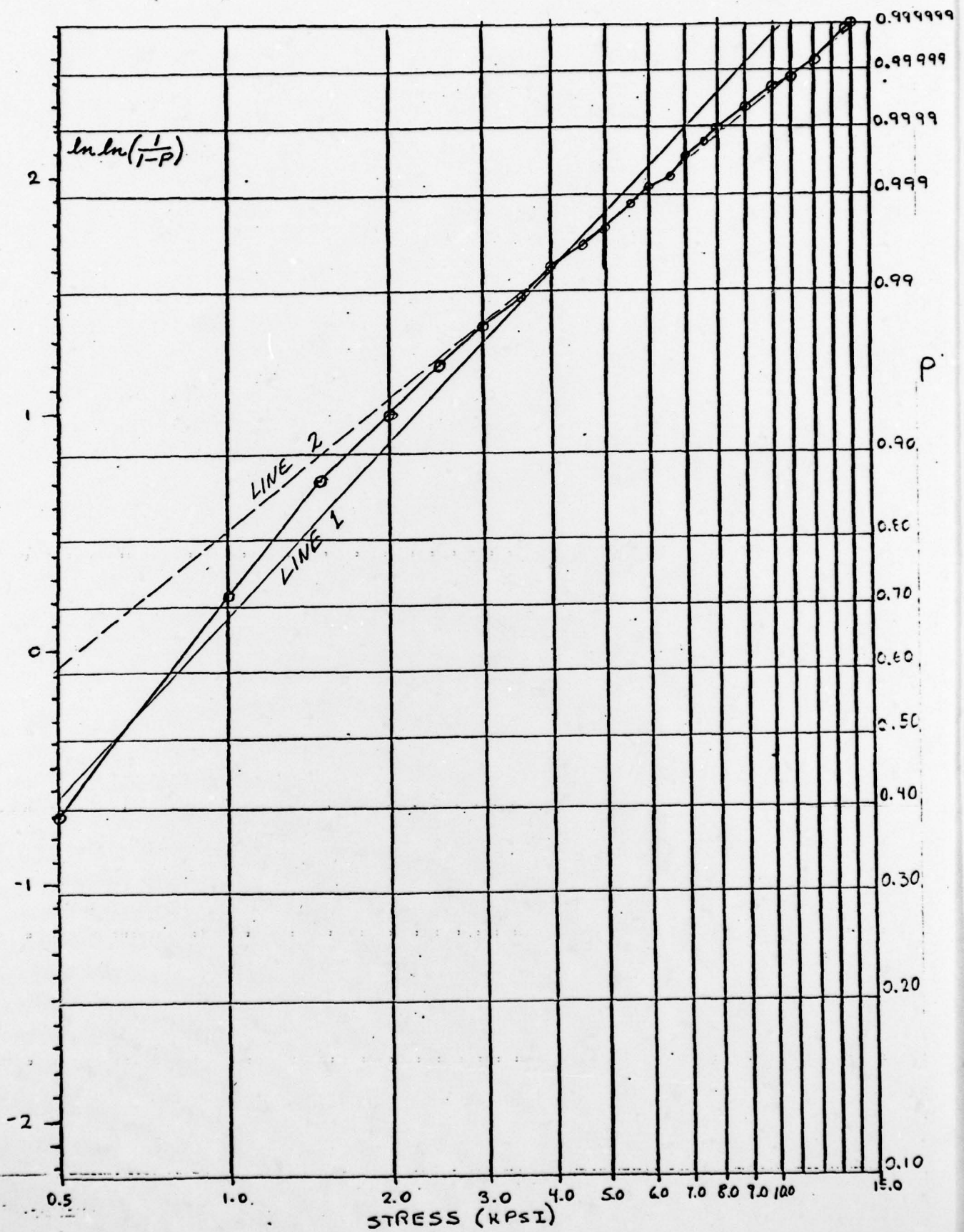
Fig. 7 : <u>Esso Malaysia</u> Stress Reversal Histogram on Weibull Probability Paper, with Possible Weibull Fits.
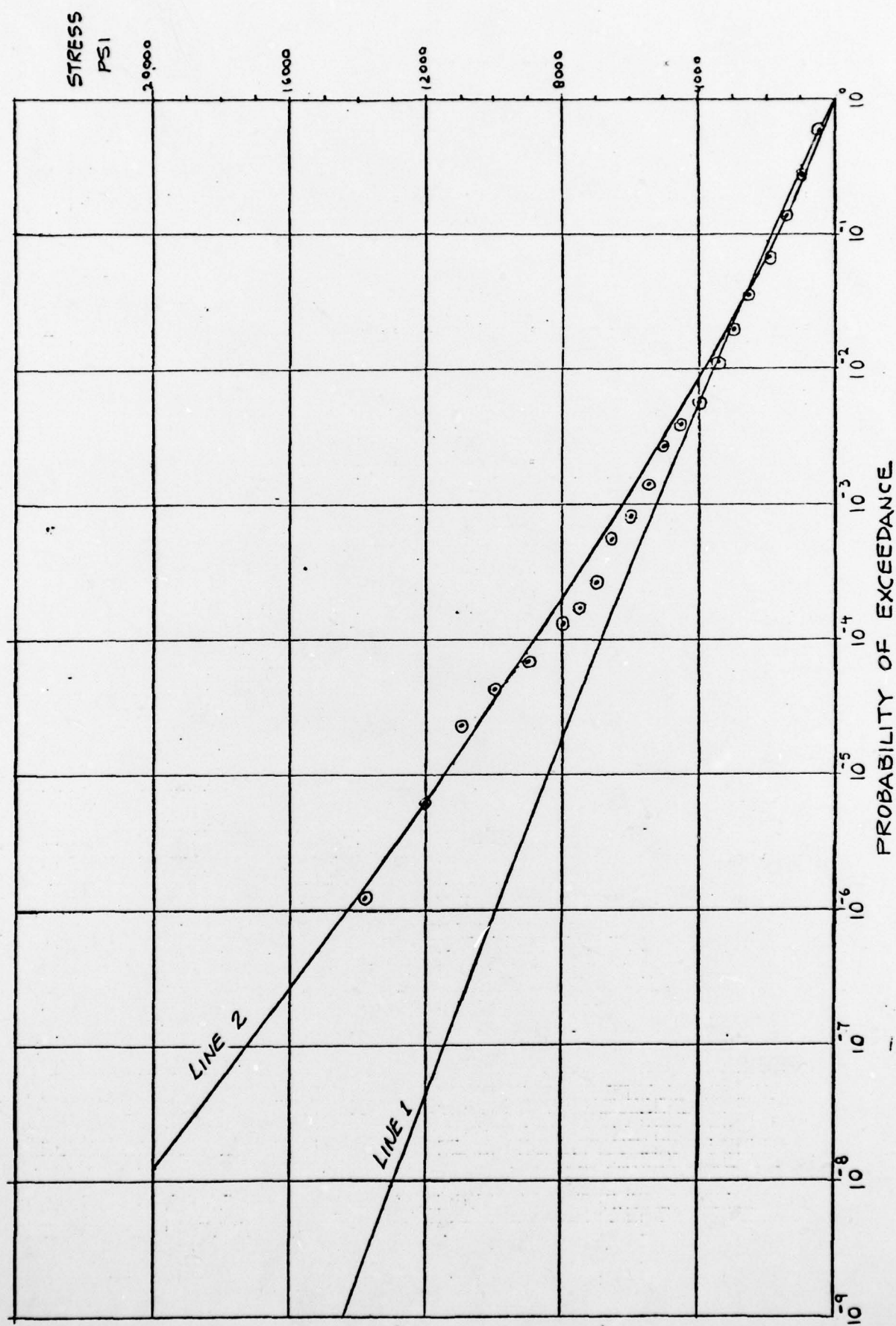
Fig. 8 : Trial Weibull Distribution Fits to Esso Malaysia
Histogram of Stress Reversals
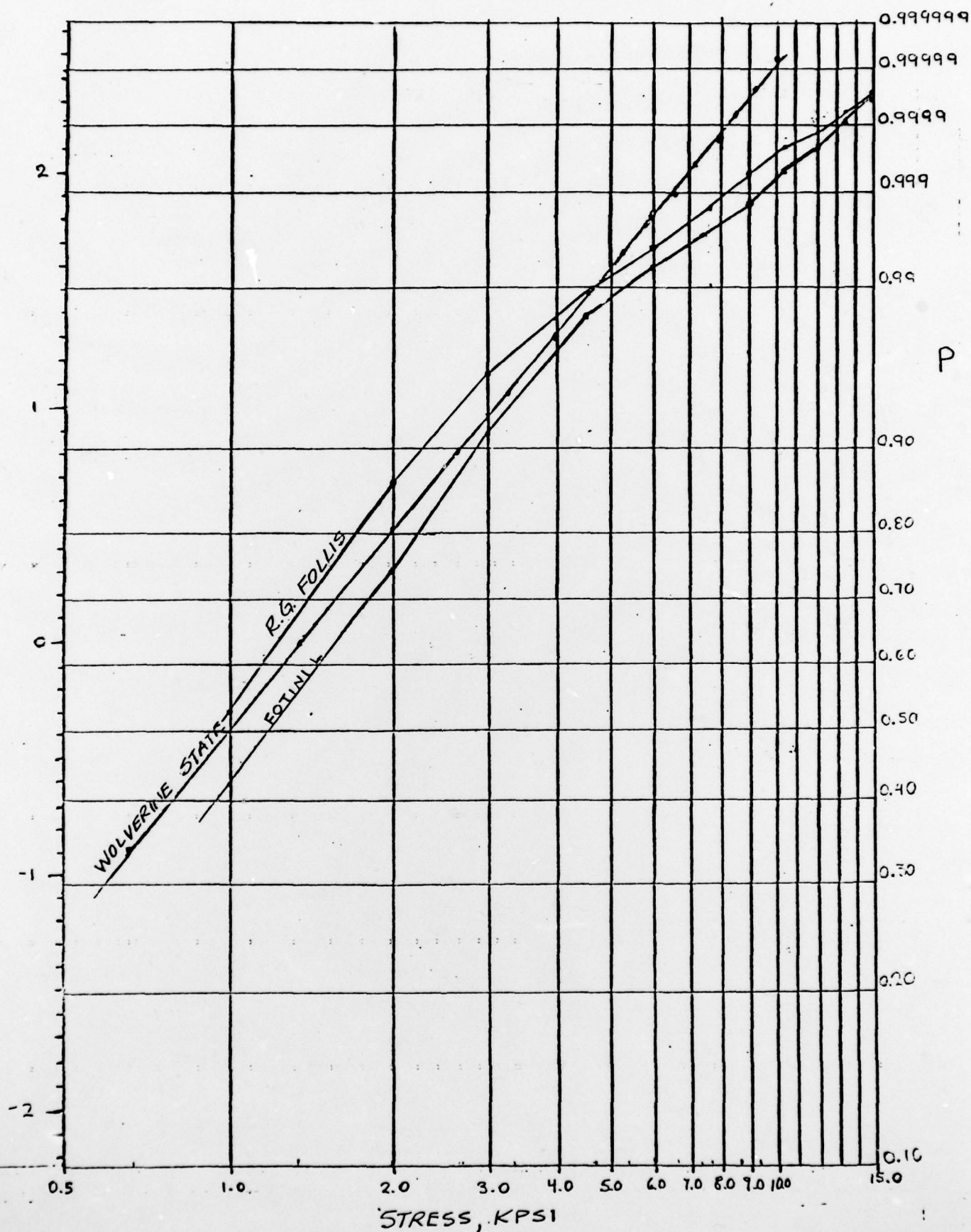
Fig. 9 : Stress Reversal Histograms on Weibull Probability Paper
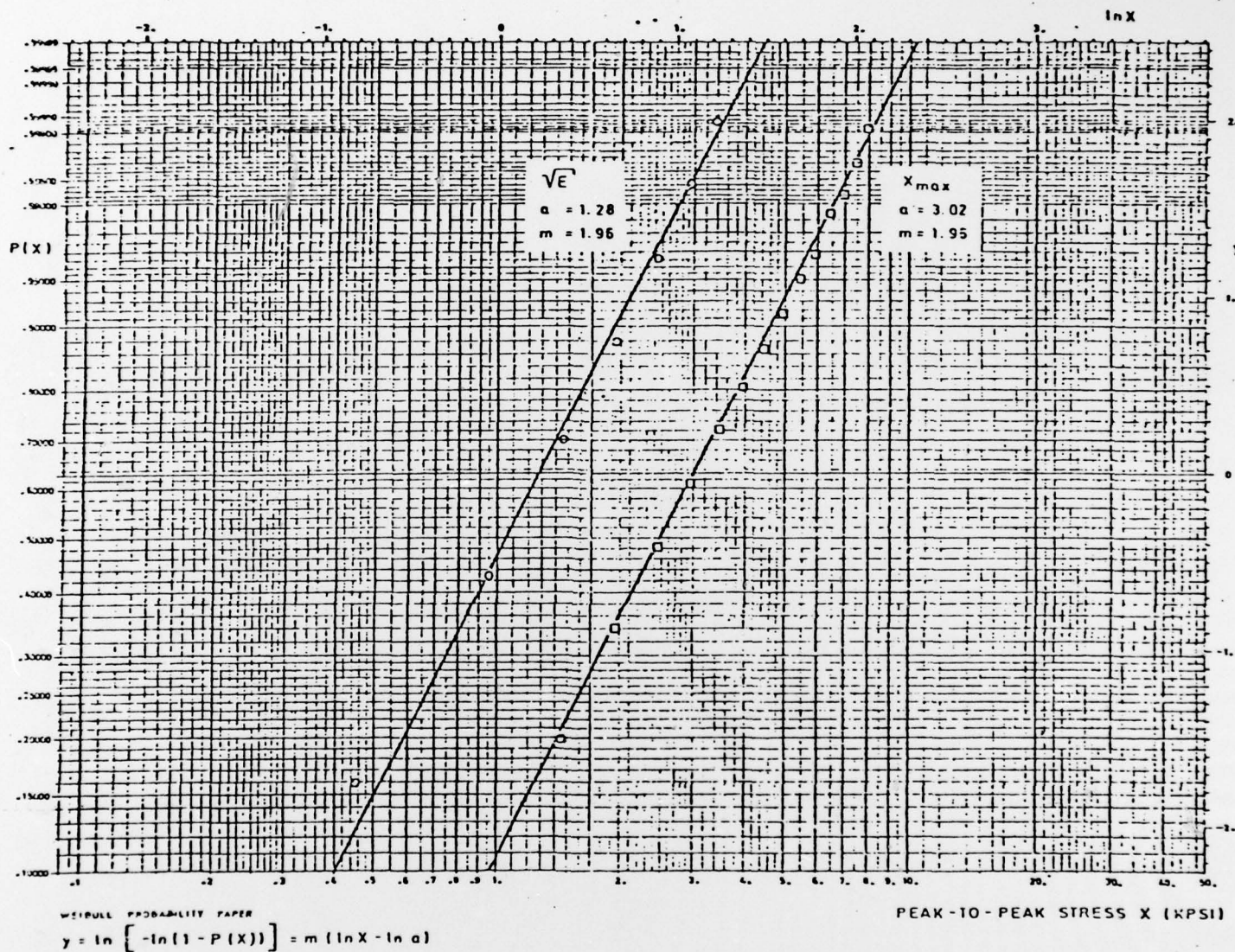from Measured Stresses on Several Ships

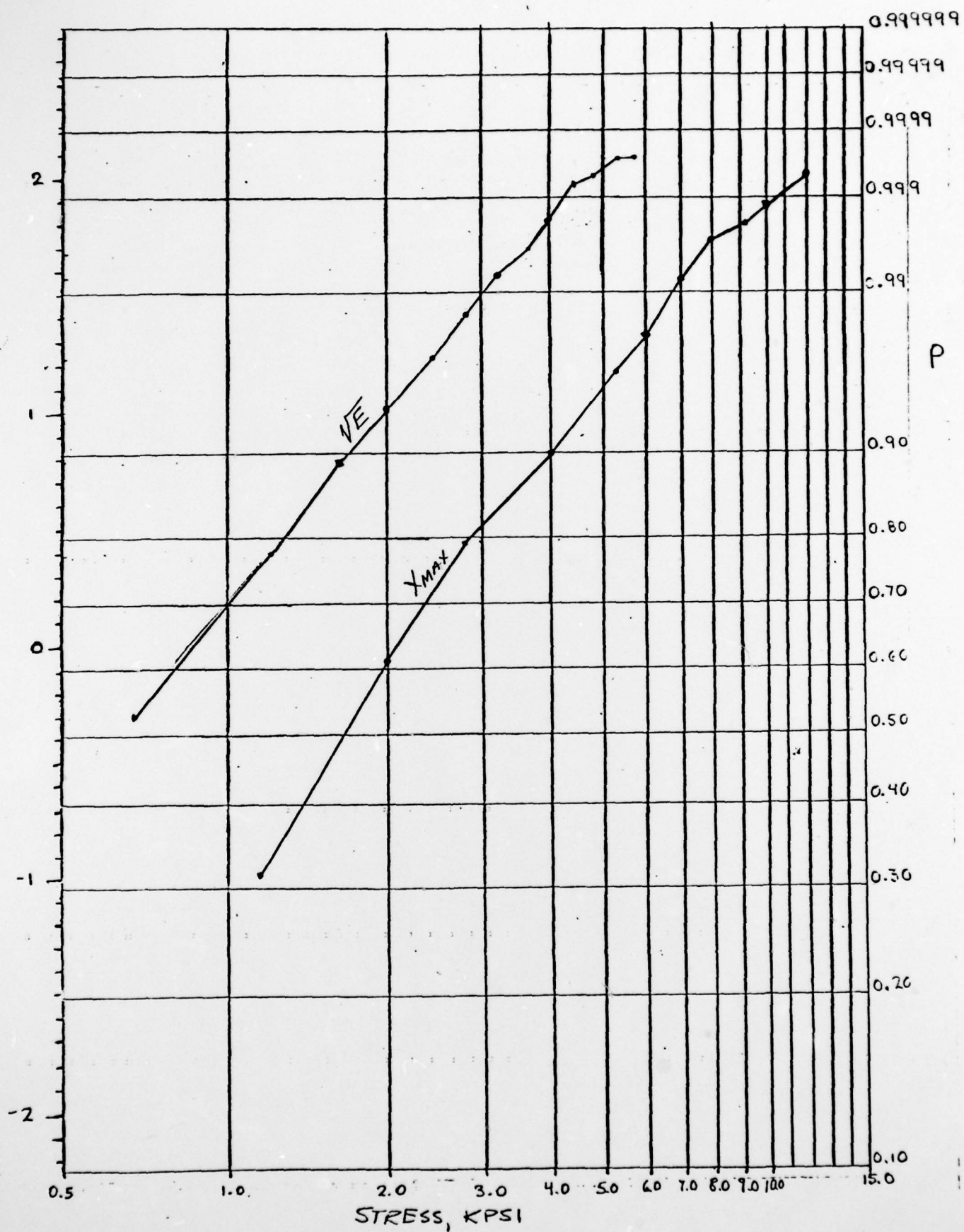Fig. 10: Weibull Predictions of $\sqrt{E}$ and X_{max} Distributions, Wolverine State (From 16)

Fig. 11: Distributions of $\sqrt{E}$ and X_{max}, <u>Esso Malaysia</u>
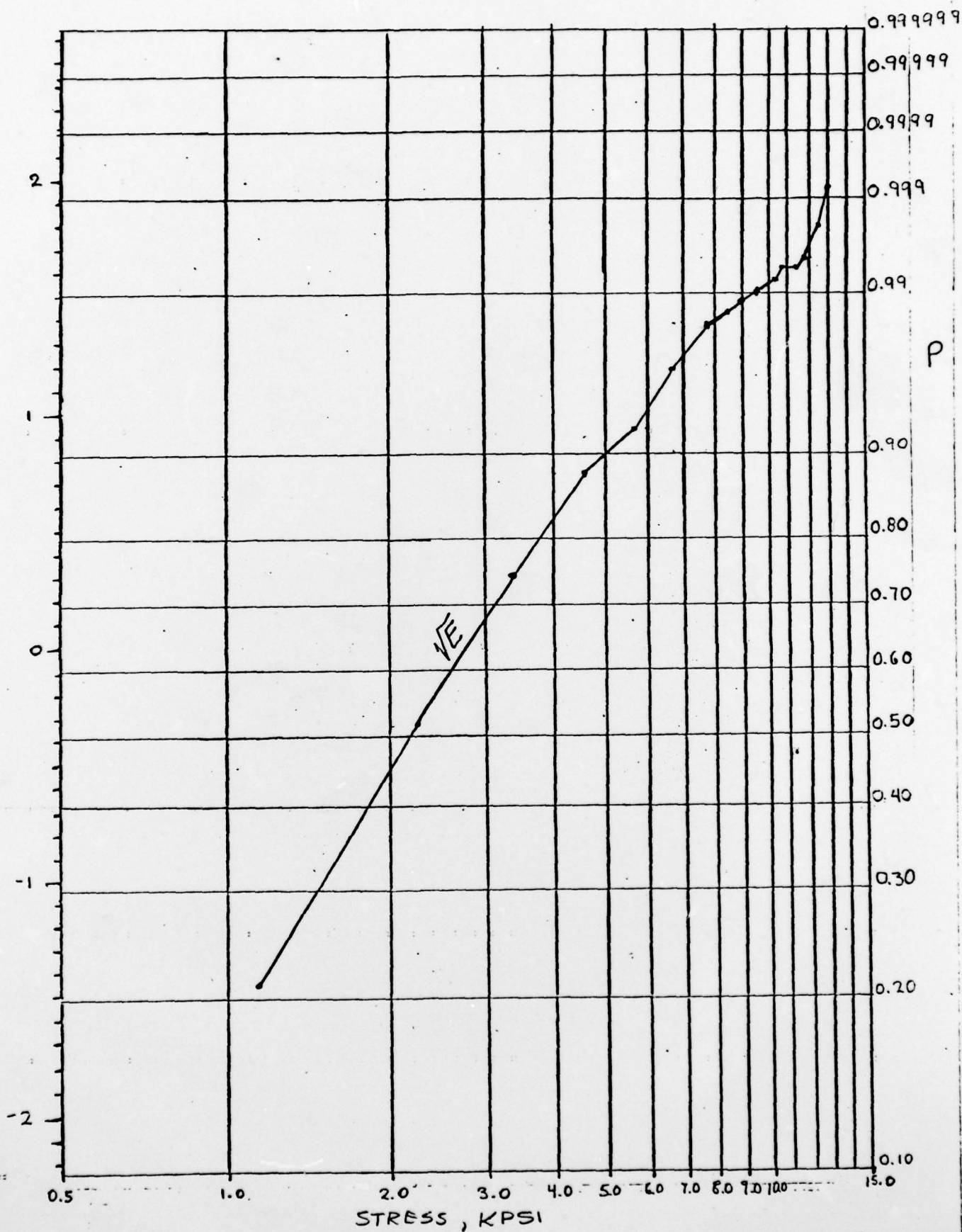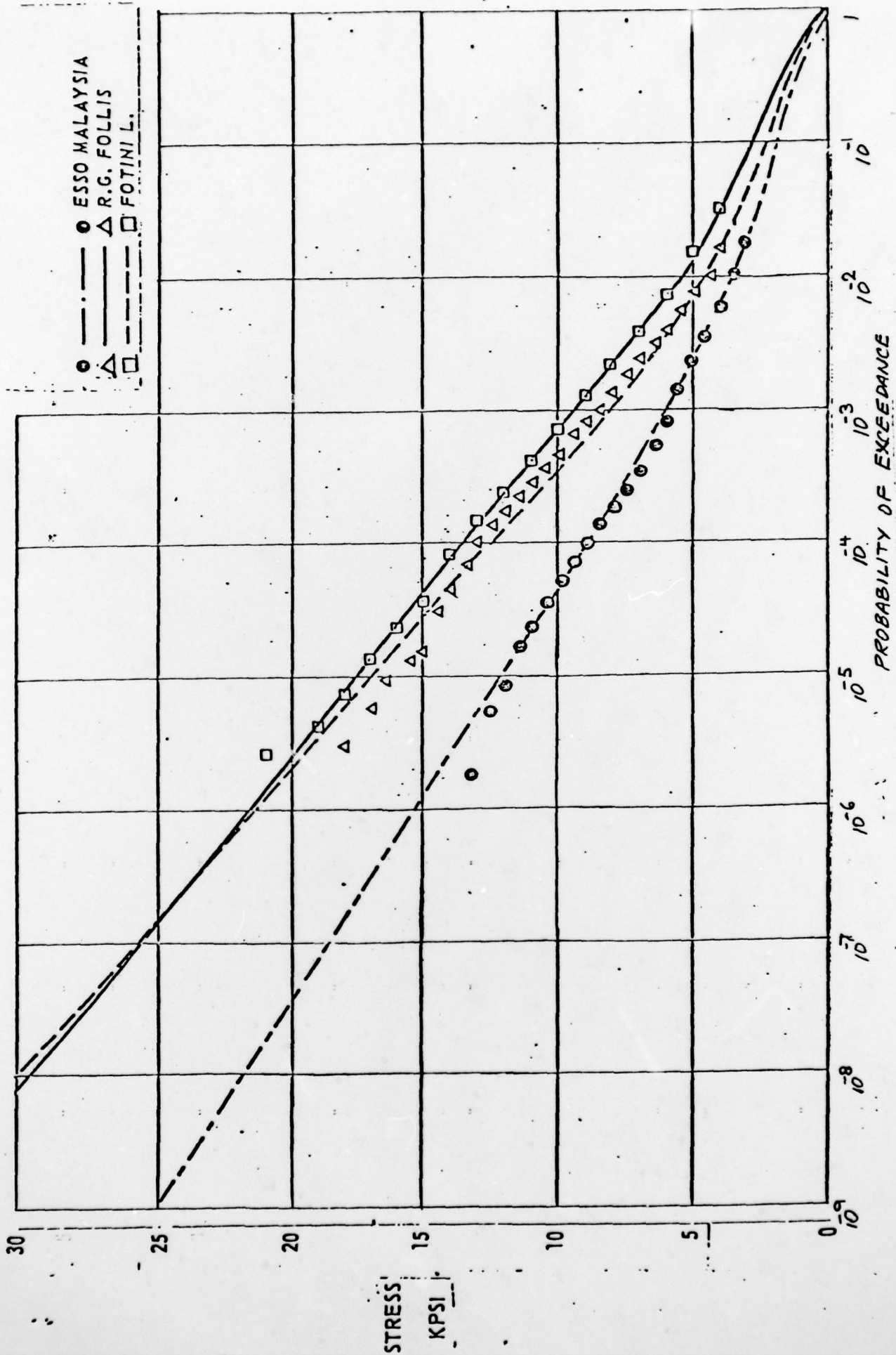
Fig. 12: Distribution of $\sqrt{E}$, <u>Sea-Land McLean</u>

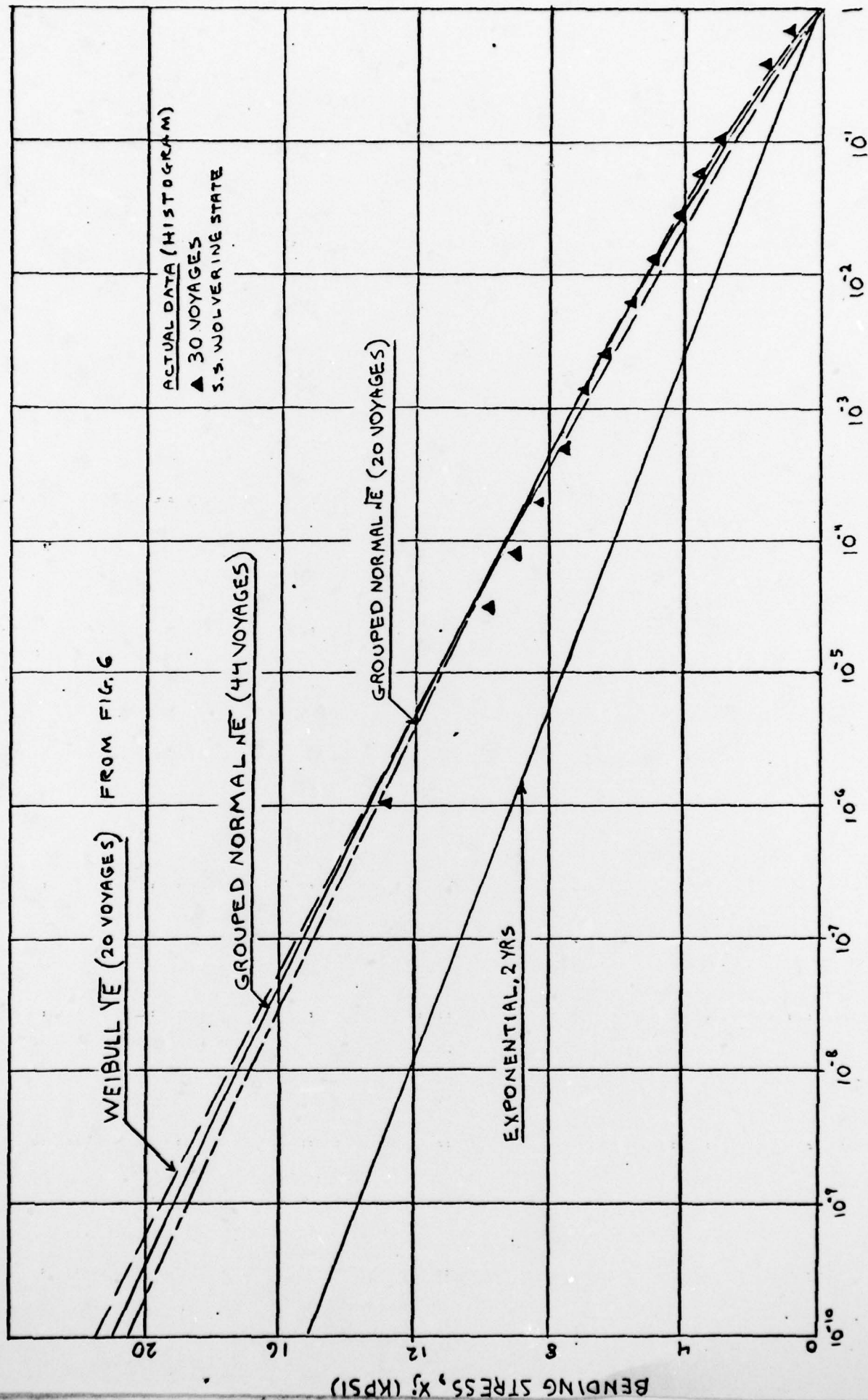Fig. 13: Long-Term Distributions by "Grouped Normal √E" Procedure Compared with Histogram Data (From 15)

PROBABILITY OF EXCEEDANCE

ACTUAL DATA (HISTOGRAM)
▲ 30 VOYAGES
S.S. WOLVERINE STATE

GROUPED NORMAL $\overline{NE}$ (20 VOYAGES)

WEIBULL $\sqrt{E}$ (20 VOYAGES) FROM FIG. 6

GROUPED NORMAL $\overline{NE}$ (44 VOYAGES)

EXPONENTIAL, 2 YRS

Q PROBABILITY OF EXCEEDANCE

BENDING STRESS, x_j (KPSI)

Fig. 14: Long-Term Predictions of Bending Stress by Various Methods,
Wolverine State (From 17 & 18)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* LONG-TERM DISTRIBUTIONS OF STOCHASTIC PROCESSES WITH APPLICATION TO SHIP HULL STRESS STATISTICS		5. TYPE OF REPORT & PERIOD COVERED Final Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(*s*) Bruce H. Stephan Robert B. Zubaly		8. CONTRACT OR GRANT NUMBER(*s*) N00014-77-C-0357
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Maritime Studies Webb Institute of Naval Architecture Crescent Beach Rd., Glen Cove, New York 11542		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-381/3-23-77 (436)
11. CONTROLLING OFFICE NAME AND ADDRESS Statistics & Probability Mathematical & Information Sciences Division, Office of Naval Research 800 N. Quincy St., Arlington, VA 22217		12. REPORT DATE December 1978
		13. NUMBER OF PAGES 72
14. MONITORING AGENCY NAME & ADDRESS(*if different from Controlling Office*) Office of Naval Research Resident Representative 715 Broadway Fifth Floor New York, NY 10003		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Distribution of this Document is Unlimited

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Stochastic processes
Probability distributions
Extreme probability distributions
Ship stress statistics

Long-term distribution of ship stresses

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

The mathematical properties of the following probability distributions and their moment generating functions are derived:

Weibull distribution;
Rayleigh distribution;
Exponential distribution;
Normal distribution;

DD $_{1 \text{ JAN } 73}^{\text{FORM}}$ 1473   EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601 |

20.

Voznesensky distribution; and
Generalized Rayleigh distribution.

The relative merits of applying these distributions to problems in
ship responses to the sea, which is described as a stationary stochastic
process, are discussed.

In Part 2, the most promising long-term distributions derived from
the above survey are applied to ship bending stress data from four ships.
It is concluded that a numerical solution of long-term distributions,
using either Weibull or Normal distributions of the short-term Rayleigh
parameters (classified by weather groups) is better than any explicit
function.